

Functional gene diversity and zone-targeted genotype prioritization using trait-diagnostic KASP markers in bread wheat (*Triticum aestivum* L.)

BINAY KUMAR SINGH

Indian Institute of Wheat and Barley Research

Ruchika Sharma

Indian Institute of Wheat and Barley Research

Nishu Raghav

Indian Institute of Wheat and Barley Research

Ravindra Singh Shekhawat

Indian Institute of Wheat and Barley Research

Sarika Jaiswal

Indian Agricultural Statistics Research Institute

M.A. Iquebal

Indian Agricultural Statistics Research Institute

O.P. Ahlawat

Indian Institute of Wheat and Barley Research

Sonia Sheoran

Indian Institute of Wheat and Barley Research

Ratan Tiwari

tiwari1964@yahoo.com

Indian Institute of Wheat and Barley Research

Research Article

Keywords: Bread wheat, Functional gene diversity, KASP assays, Population structure, Marker-assisted selection, Zone-targeted prioritization

Posted Date: February 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8710524/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Functional gene diversity is essential for adaptation and sustained genetic improvement in bread wheat. In this study, 40 KASP assays targeting 40 functional SNPs across 35 genes linked to adaptation, stress resilience, quality, and yield were used to quantify allelic variation and develop a strategy for region-specific genotype deployment. Functional diversity and population structure were assessed in 178 wheat genotypes, including genetic stocks (80), released varieties (72), and landraces (26). High-throughput KASP assays classified all loci as favourable or unfavourable. Phylogenetic, principal coordinate, and Bayesian STRUCTURE analyses consistently identified two major subpopulations: a founder-centric group fixed for productivity- and disease-resistance alleles, and a recombination-driven group enriched for adaptive and stress-resilience alleles. Among trait categories, adaptation-related loci showed the highest mean gene diversity ($H_e = 0.420$), while loci linked to yield and quality had lower polymorphism, indicating strong directional selection during breeding. Landraces had the highest proportion of rare favourable alleles (mean = 9.35), highlighting their importance as reservoirs of underutilized functional variation for pre-breeding. The complementary distribution of alleles between founder- and recombination-derived groups reflected the combined effects of historical selection and introgression on the genomic structure of Indian wheat. Using this functional genomic dataset, a rule-based, zone-weighted suitability scoring model was developed to integrate allele states with breeding priorities across India's five wheat production zones. This model enabled quantitative, zone-specific genotype prioritization and identification of regionally adapted donors with complementary adaptive alleles. The strategy turned functional markers into practical tools for strategic allele deployment and established a scalable decision-support system for precision breeding.

1. Introduction

Bread wheat (*Triticum aestivum* L.) is an annual, predominantly autogamous species in the Triticeae tribe of the grass family (Poaceae). It is an allohexaploid species with 21 chromosome pairs organized into three subgenomes (A, B, and D), with a genome composition of BBAADD ($2n = 6x = 42$) [1]. Bread wheat has evolved through two rounds of allopolyploidization, resulting from cycles of 'diverge-and-merge' among its progenitors. The first event, approximately 800,000 years ago, led to the formation of tetraploid wild emmer wheat (*T. turgidum* ssp. *dicoccoides*) through hybridization between two diploid species: the male donor of the A subgenome, related to *T. urartu*, and the female donor of the B subgenome, related to *Aegilops speltoides* [2, 3]. The second allopolyploidization event, occurring approximately 9,000 years ago, involved allotetraploid wheat (*T. turgidum*; BBAA) as the female donor and *Aegilops tauschii* (formerly *Aegilops squarrosa*; DD) as the male donor, producing hexaploid bread wheat [2]. The D genome, derived approximately 5.37 million years ago from the D lineage, originated from hybridization between the ancestral donors of the A and B subgenomes [2]. Following its domestication in the Fertile Crescent, the cultivation of bread wheat spread worldwide, predominantly in temperate regions but also in subtropical and tropical habitats [4]. By 2024, global wheat production reached a record 791.24 million tonnes, with China and India as the leading producers, producing 136.59 and 110.55 million tonnes, respectively [5].

Wheat cultivation in India dates back to 3000 BC [6], with farmers traditionally growing heterogeneous landraces and mixtures known as "sorts" until the early 1900s [7]. Systematic wheat improvement began in 1905, focusing on enhancing the yield of native tall landraces. These efforts led to the development of several popular wheat varieties, such as NP4, NP6, and NP12, through pure-line selection from indigenous materials. These varieties were later used in hybridization programmes to develop higher-yielding varieties such as C591 [7]. Exotic germplasm played a critical role in subsequent wheat improvement programs, especially in resistance breeding, leading to the development of rust-resistant varieties such as NP783 and NP784 (resistant to brown rust), NP785 and NP786 (resistant to yellow rust), NP789 and NP790 (resistant to black rust), and NP809 (resistant to all three rusts and loose smut) [7]. Concurrently, distant hybridization produced varieties such as C286 and Niphad 4 [6]. However, due to their tall, thin straw, almost all cultivated varieties exhibited poor responsiveness to fertilizers or irrigation [8].

The breakthrough in yield improvement came in 1963 with the introduction of Mexican semi-dwarf segregating material carrying the *Norin* dwarfing genes. This led to the development of amber-grained, high-yielding, semi-dwarf varieties like Kalyansona and Sonalika [9]. Hybridization between local and exotic dwarf wheat, followed by pedigree selection, led to the development of popular varieties such as HD 2329, thereby further boosting wheat production in India. The introduction of 'Veery' lines developed at CIMMYT, Mexico [10], carrying the 1BL/1RS rye translocation, led to the successful development of varieties such as PBW 343, marking another significant milestone in Indian wheat improvement [11].

The All India Coordinated Wheat Improvement Project (AICWIP), operational since 1965, has played a critical role in evaluating high-yielding wheat varieties across India. AICWIP facilitates access to germplasm from international collaborators, including CIMMYT and ICARDA. Since its inception, 546 high-yielding wheat varieties have been released for diverse agroclimatic zones, increasing productivity from 913 kg/ha in 1964-65 to 3,615 kg/ha in 2023–2024 [12]. However, wheat breeding continues to face challenges in India due to global climate change, decreasing arable land, and water shortages in arid and semi-arid regions.

Polyploidization [2], domestication [13], gene flow [14, 15], and post-domestication selection [16] have significantly influenced the genetic structure of bread wheat. Nonbrittle rachis (*Btr1*, *Btr2*) and free-threshing (*Q*, *Tg*) [17] were among the earliest traits selected for wheat domestication. Traits such as vernalization (*Vrn-A1*, *Vrn-B1*, *Vrn-D1*) [18, 19, 20] and photoperiod sensitivity (*Ppd-D1*) [21] were selected to enhance wheat adaptation to diverse environments. The selection for reduced plant height (*Rht-B1*, *Rht-D1*) [22] increased input-use efficiency and prevented lodging. Modern wheat breeding efforts pay special attention to select genes related to pre-harvest sprouting (*PHS1*, *Sdr-B1*, *MFT*, *VP-1B*, *PH*) [23, 24, 25, 26, 27, 28] and yield-related traits (*Sus1-7A*, *Sus1-7B*, *Sus2-2A*, *Sus2-2B*, *GW2-6A*, *GW2-6B*, *GS-D1*, *TaSBEIII-A*, *GASR-A1*, *MOC-7A*) [29, 30, 31, 32, 33, 34, 35, 36], which influence grain size, weight, and yield traits. Glutenin subunits (*Glu-A1*, *Glu-B1*, *Glu-D1*) [37, 38], kernel hardness (*Pina-D1*, *Pinb-D1*, *Pinb2-B2*) [39, 40], starch content (*Wx-B1*) [41], protein content (*NAM-B1*) [42], and key enzymes such as cell wall invertase (*Cwi-4A*) [43] and IAA-glucose hydrolase (*TGW-6*) [44] are crucial targets for improving suitability for diverse end uses. Additionally, significant attention is given to biotic and abiotic stress-related genes, including drought tolerance (*Dreb B1*, *1-feh-w3*) [45, 46], Fusarium head blight resistance (*Fhb1*, *PFT*) [47, 48], leaf rust resistance (*Lr14a*, *Lr34*, *Lr68*) [49, 50, 51], stripe rust resistance (*Yr15_R8*, *Yr15_R5*) [52], and the 1BL/1RS translocation [53].

Marker-assisted selection has become a powerful tool in modern wheat breeding, especially for traits controlled by mutations with significant and predictable genetic effects. Advances in high-throughput functional genotyping, particularly Kompetitive Allele-Specific PCR (KASP) assays that target causative SNPs, now enable precise differentiation between favourable and unfavourable alleles at trait-diagnostic loci [54, 55, 56]. Because functional markers focus on polymorphisms within the causal gene sequence, their predictive accuracy is minimally affected by linkage disequilibrium with nearby loci [57], making them especially suitable for marker-assisted selection. Although over 150 functional markers linked to more than 100 adaptive, stress-resilience, quality, and yield-related loci have been converted into reliable KASP assays [56], their systematic use to assess functional gene diversity, analyze population structure, identify complementary allele pools, and guide zone-specific breeding decisions remains limited in Indian wheat.

Given India's diverse production ecologies, ranging from the intensively managed North-Western Plains to the drought-prone Central and Peninsular Zones and the cool, high-altitude hill regions, aligning adaptive alleles with zone-specific breeding priorities is crucial for maximizing genotype–environment fit. Integrating functional-allele data with zone-level trait requirements provides a biologically informed and logical system to improve selection accuracy, preserve genetic diversity, and speed up durable genetic gain. In this context, the present study examines allelic diversity across 40 functional SNPs spanning 35 genes associated with adaptation, stress tolerance, grain quality, and yield in a panel of 178 Indian bread wheat genotypes representing genetic stocks, released cultivars, and landraces. Using high-resolution KASP genotyping, phylogenetic and STRUCTURE analyses, and a rule-based, zone-weighted suitability scoring model, we (i) characterized functional diversity and identify signatures of historical selection, (ii) defined founder-derived and recombination-driven genetic groups, (iii) measured rare-allele enrichment in landraces, and (iv) created a novel operational system for zone-targeted genotype prioritization. By directly linking functional genomic variation to environment-specific breeding strategies, this study establishes an integrated platform for deploying adaptive, high-value alleles to improve productivity, resilience, and quality in Indian wheat improvement programs.

2. Materials and methods

2.1. Plant materials and DNA extraction

A diverse panel of 178 bread wheat (*Triticum aestivum* L.) genotypes was assembled to assess functional genetic diversity across various germplasm categories. This panel included 80 genetic stocks, primarily wild introgression lines developed for traits related to biotic and abiotic stress tolerance and biological nitrification inhibition (BNI), along with 72 released varieties and 26 landraces (Table 1). All materials were sourced from the Germplasm Resource Unit of the ICAR–Indian Institute of Wheat and Barley Research (ICAR-IIWBR), Karnal, Haryana, India, the national centre for wheat germplasm conservation and distribution. Plants were grown under controlled conditions, and young, healthy leaf tissues were collected at the seedling stage for genomic DNA extraction. DNA was isolated using the cetyltrimethylammonium bromide (CTAB) method described by Murray and Thompson [58], with slight modifications to enhance yield and purity. The concentration and integrity of the extracted DNA were verified with agarose gel electrophoresis and a NanoDrop™ spectrophotometer, and only high-quality DNA samples proceeded to further analysis.

Table 1
Details of plant materials used in the study.

S. No.	Accession	Pedigree	Accession type
1	KBRL22	HD29/W485	Genetic stock
2	MP1201	CMH-82-A-1294/2*KAUZ//MUNIA/CHORLITO/3/MILAN	Variety
3	HD3118	ATTILA*2/PBW-65//WEEBILL-1*2/TUKURU	Variety
4	AKAW3717	HW2035/NI5439	Genetic stock
5	HD2932	KAUZ/STAR//HD2643	Variety
6	FLW3	UP2338/China 84-40022	Genetic stock
7	HS240	AVOCA/KALYANSONA/BLUEBIRD/WOODPECKER(SIB)/(SIB)PAVON-76	Variety
8	AKW2862-1	VEERI-6//DOVE(SIB)/(SIB)BUCKBUCK	Genetic stock
9	DBW31	DBW14/PBW343	Genetic stock
10	RAJ3765	HD2402/VL639	Variety
11	HW2001	SONALIKA//TR-380-14*7/3-AG-14/3/5*SONALIKA	Variety
12	KRL19	PBW255/KRL1-4	Variety
13	HPW349	OASIS/KAUZ//4*BACANORA-88/3/PASTOR/4/KAUZ*2/YACO//KAUZ	Variety
14	HD2687	CPAN-2009/HD-2329	Variety
15	DBW288	SOKOLL/3/PASTOR//HXL7573/2*BAU/4/GLADIUS	Genetic stock
16	PUSA2019A	CS CMS*7/HW2019	Genetic stock
17	PBW703	BW9250*3/YR10/6*AVOCET/3/BW9250*3//YR15/6*AVOCET	Genetic stock
18	D2	<i>Ae. tauschii</i> (13781)/4*HD2967	Genetic stock
19	PBW771	BW3246/2*DBW17	Variety
20	HUW12	NP876/CIANO 66	Variety
21	HD2864	DL509-2/DL377-8	Variety
22	FLW30	PBW343/Yr15//PBW34/LR28	Genetic Stock
23	IC321866	SK 21, BROWN SPIKE LOCAL WHEAT, RAISEN, MP	Landrace
24	H567.71	Developed through a cross involving IAS20 with pedigree <i>COLONIAS//FRONTANA/KENYA58</i>	Genetic stock
25	CMH77.308	H567.71/3*PELOTAS-ARTHUR	Genetic stock
26	LGM205	LOCAL COLLECTION FROM GUJARAT	Landrace
27	SEL111	CHINESE-SPRING-monosomic-5B/(SE.CE)RUSSIAN-RYE/C-591//SONALIKA	Genetic stock
28	HW2049	HD-2285*6//SUNSTAR*6/C-80-1	Genetic Stock
29	SONORA64	YAKTANA-54//NORIN-10/BREVOR/3/2*YAQUI-54	Variety
30	D18	NIDW15/CS/ <i>Ae. speltoides</i>	Genetic stock
31	PWR4099	CBHW-CHN-89-R-42940-CHN-S-2BV-97-(EC-4141480)	Genetic Stock
32	UAS320	UAS-257//GW-322/DWR-195	Genetic Stock

S. No.	Accession	Pedigree	Accession type
33	PBW698	BW9250*3/YR10/6*AVOCET/3/BW9250*3//YR15/6*AVOCET	Genetic stock
34	BNI NOVOJOA	CSMONO3B/3/CS/LE.RA//CS/4/5*NAVOJOA	Genetic stock
35	NW4035	WEAWER/WL-2926//SW-89-3064	Genetic stock
36	PBW644	PBW-175/HD-2643	Variety
37	LOK44	CPAN4041/LOK1	Genetic stock
38	D55	PBW 698/(CS/ <i>Ae speltoides</i>)/WH 1105	Genetic stock
39	BNI MUNAL	CSMONO3B//CS*2/LE RA/3/CS PH/4/6*MUNAL	Genetic stock
40	KALYANSONA	PENJAMO-62(SIB)/GABO-55	Variety
41	WH595	PRL6045/NAC76	Genetic stock
42	HS526	CMH82A.1294/2*KAUZ//MUNIA/CHTO/3/MILAN	Genetic stock
43	PBW596	PBW-343/DHARWAD-DRY//PBW-343	Variety
44	RNB1001	HD2687/MACS2846	Genetic stock
45	UP262	SONALIKA(S-308)/BAJIO-66	Variety
46	IC296432	WL711*4/ <i>Aegilops geniculata</i> Acc. 3547	Genetic stock
47	HTW6	HEAT TOLERANT STOCK	Genetic stock
48	KHARCHIA65	KHARCHIA LOCAL/EG 953	Variety
49	NP846	NP760/RIONEGRO	Variety
50	IC321856	SK11, PISSILocal FROM GUNA, MP	Landrace
51	IC321849	SK4, LOCAL COLLECTION FROM RAISEN, MP	Landrace
52	SONALIKA	II-53-388/ANDES//(SIB)PITIC-62/3/LERMA-ROJO-64	Variety
53	IC321910	SK65, SHARBATI LOCAL FROM AJGAH	Landrace
54	HW5207-1	HW3029/ <i>Yr15</i>	Genetic stock
55	HD3043	PAJONAL/BOBWHITE//OPATA-85*2/3/CROC-1/TR.TA(224)//OPATA-85	Variety
56	HTW11	HEAT TOLERANT STOCK	Genetic stock
57	BNI SONORA	Developed by integrating BNI trait from <i>Leymus racemosus</i> into SONORA	Genetic stock
58	WL5634	KAVKAZ-CNO-INIA/WL903	Genetic stock
59	MBL5	NP4/HB209	Genetic stock
60	FLW29	PBW343/CAPPELLE DESPREZ/PBW343/Lr28(CS2D2M3/8)F7	Genetic stock
61	HINDI62	PBW-226/HW-1042//HD-2285	Genetic stock
62	DBW46	PBW343/INQ21	Genetic stock
63	GW322	PBW173/GW 196	Variety
64	WBM1591	PYN/BAU//MILAN	Genetic stock

S. No.	Accession	Pedigree	Accession type
65	K307	K8321/UP2003	Variety
66	DBW303	WBLL1*2/BRAMBLING/4/BABAX/LR42//BABAX*2/3/SHAMA*2/5/PBW343*2/KUKUNA*2/FRTL/PIFED	Variety
67	D16	PBW698/(<i>T. turgidum</i> cv Neodur/ <i>Th bessarabicum</i>) EC- 787010	Genetic stock
68	GW2002-18	DL51-91/GW265	Genetic stock
69	HW4005	KALYANSONA*6//C 86 – 8/KALYANSONA F4	Variety
70	FLW5	UP2338/ARKAN	Genetic stock
71	HD2888	C 306/ <i>T. sphaerococcum</i> //HW 2004	Variety
72	CBW38	CNDO/R 143/ENTE/MEXI_2/3/Ae.SQUARRO SA (TAUS)/4/WEAVER/5/2*PASTOR	Variety
73	GW366	DL802-3/GW232	Variety
74	BACANORA88	JUP/BJY//URES	Variety
75	DBW51	SITELLA/MILAN	Genetic stock
76	KRL1-4	KHARCHIA65/WL711	Variety
77	ISD215	KS/ <i>T. turgidum</i> //h38/3/SKA// <i>T. carthlicum</i> //33/SKA/BLUEBIRD//HD2122	Genetic stock
78	WH1061	WEAVER/WL 2926//SW 89.3064	Genetic Stock
79	HD3086	DBW14/HD2733//HUW468	Variety
80	MACS6222	HD2189*2//MACS2496	Variety
81	BNI ROELFS	CSMONO3B//CS*2/LE RA/3/CS PH/4/5*ROLF07	Genetic Stock
82	KRL35	HD-2160/KRL-1-4 (KHARCHIA-LOCAL/WL-711)	Genetic Stock
83	FLW11	WH 542*/Hobbit	Genetic stock
84	FLW13	WH542*/Yr15(CN25087)	Genetic stock
85	HI1500	HW2002*2//STREMPALLI/PNC 5	Variety
86	HD3016	PBW-65/2*PASTOR	Genetic stock
87	K1213	HD2733/HD2285	Variety
88	RAJ4356	NOT AVAILABLE	Genetic stock
89	BNI DBW222	Developed by integrating BNI trait from <i>Leymus racemosus</i> into DBW222	Genetic stock
90	D14	DBW39 /(Highbury/ <i>Ae mutica</i>)	Genetic stock
91	LOK1	SONALIKA/CHOTI-LERMA	Variety
92	WH542	JUPATECO-73/BLUEJAY//URES-81	Variety
93	DBW327	NELOKI//SOKOLL /EXCALIBUR	Variety
94	D127	SYN224/3*WH1105	Genetic stock
95	ALDAN	MAYA-74(SIB)/4/KLEIN-RENDIDOR/2*SONORA-64//INIA-F-66/3/CIANO-F-67/5/NEWTATCH//STRAMPELLI/YAQUI-50	Genetic stock
96	GL23	LOCAL COLLECTION FROM GUJARAT	Landrace
97	YR10	Avocet+Yr10	Genetic stock
98	YR15	Avocet+Yr15	Genetic

S. No.	Accession	Pedigree	Accession type
			stock
99	IC322007A	SK 162, LOCAL WHEAT FROM CHAMOLI	Landrace
100	NB29-13-5-9	LOCAL COLLECTION FROM GUJARAT	Landrace
101	MP3288	DOVE/BUC/DL788-2	Variety
102	PBW723	PBW343 + Lr57/Yr40 + Yr15	Variety
103	D22	DBW39/(Highbury/ <i>Ae mutica</i>)/DBW 39	Genetic stock
104	WL711	HD1982/K816	Variety
105	HI1579	KAUZ//ALTAR84/AOS/3/MILAN/KAUZ/4/HUITES	Genetic stock
106	DBW316	DBW18/DBW 66	Variety
107	DL1266-5	SFW/VAISHALI	Genetic Stock
108	FKW1	UP2338*4/CHINA84-40022	Genetic Stock
109	DBW252	PFAU/MILAN/5/C. HEN/AEGILOPS SQUARROSA (TAUS)//BCN/3/VEE#7/BOW/4/PASTOR	Variety
110	DBW39	ATTILA/HUI	Variety
111	CPAN3045	COLLECTION FROM PUNJAB	Landrace
112	IC322001	SK 156,VIKAS GEHUN LOCAL WHEAT FROM GOPESHWAR	Landrace
113	HD2967	ALONDRA/CUCKOO//URES-81/HD-2160-M/HD-2278	Variety
114	DPW621-50	KAUZ//ALTAR-84/(AOS)AWNED-ONAS/3/MILAN/KAUZ/4/HUITES	Variety
115	DBW377	NADI#2*2/6/BECARD#1/5/KIRITATI/4/2*SERI.1B*2/3/KAUZ*2/BOW//KAUZ	Variety
116	HD30	HD-2160/HD-1977//HD-7949/HD-1944/HD-2136	Genetic stock
117	IC321979	SK 134, LATE LOCAL VARIANT FROM NAGNI, CHAMBA	Landrace
118	HD29	HD2160-HD1977/HD7449-HD 1944/HD2136	Genetic stock
119	DBW50	KAUZ//ALTAR84/AOS/3/MILAN/KAUZ/4/HUITES	Variety
120	PBW827	PBW677/PBW703	Variety
121	PBW660	WG6761/WG6798	Variety
122	RIL1	DHARWAD DRY/DPW621-50	Genetic stock
123	DTS16	HD3070/C306	Genetic stock
124	DBW187	NAC/TH.AC//3*P VN/3/MIRLO/BUC /4/2*PASTOR/5/K ACHU/6/KACHU	Variety
125	VL900	KAUZ/3/MYNA/BUL/FLK/4/MILAN	Genetic stock
126	VL798	CPAN-3004/CPAN-1922/PBW-65	Genetic stock
127	HI1580	HD 2402/HW3007	Variety
128	LGM133	LOCAL COLLECTION FROM GUJARAT	Landrace
129	W485	HD1982/WH533	Genetic stock
130	SB14	BH1146/HD2967	Genetic stock
131	BNI DBW187	Developed by integrating BNI trait from <i>Leymus racemosus</i> into DBW187	Genetic stock
132	IC321950	SK 105, IMPROVED VARIETY FROM DOLOIGAON, ASSAM	Landrace
133	IC321946	SK 101, DWARF WHEAT FROM MORAGAON, ASSAM	Landrace

S. No.	Accession	Pedigree	Accession type
134	NW1014	HAHN'S'	Variety
135	VL639	VL421/CPAN1535	Genetic stock
136	PBW674	PBW441/PBW476	Variety
137	DBW14	RAJ3765/PBW343	Variety
138	IC321984	SK 139, BOLD GRAIN OLD VARIETY FROM TEHRI, UTTRANCHAL	Landrace
139	DBW371	BORL14/CHIPAK	Variety
140	HP1531	COLLECTION FROM PUNJAB	Landrace
141	LGM139	LOCAL COLLECTION FROM GUJARAT	Landrace
142	KBRL10	HD29/HP1531	Genetic stock
143	NARMADA4	GB-AVS/N14/3/PW5//TH/NP165	Variety
144	IC212162	COLLECTED FROM JAMNAGAR	Landrace
145	IC321998	SK 153, PAHARI LOCAL FROM GOPESHWAR	Landrace
146	COW(W)1	HD2646//HW2002A/CPAN3057	Variety
147	CHIRYA3	CS/AG.CU//GLEN/3/ALD/PVN/4/NINGMAI NO.4/OLESON//ALD/YANGMAI NO.4	Variety
148	RIL110	DHARWAD DRY/DPW621-50	Genetic stock
149	DTS15	GW366/SQ1	Genetic stock
150	GL6	LOCAL COLLECTION FROM GUJARAT	Landrace
151	DHARWADDRY	DWR39/C306//HD2189	Variety
152	DTS1	DBW39/SQ1	Genetic stock
153	KINGBIRD	TAM200/TUI//PIHA	Variety
154	DBW296	SOKOLL/3/PASTO R//HXL7573/2*B AU/4/MASSIV/PP R47.89C	Variety
155	DTS8	DBW39/SQ1	Genetic stock
156	K9351	K72/K8027//K72	Variety
157	HD2009	LERMA-ROJO-64-A/NAINARI-60	Variety
158	DBW372	FD08114/BECAR D#1//BOKOTA	Variety
159	HD2329	HD1962/E4870/3/K65/5/HD1553/4/UP262	Variety
160	IC212161	COLLECTED FROM JAMNAGAR	Landrace
161	UP2303	HD2296/ska	Variety
162	LGM132	LOCAL COLLECTION FROM GUJARAT	Landrace
163	DBW273	FRANCOLIN #1*2//ND 643/2* WBLLI	Variety
164	DBW88	KAUZ//ALTAR-84/AWNED-ONAS/3/MILAN/KAUZ,MEX/4/HUITES	Variety
165	DBW332	MUTUS/ROLF07/ /MUCUY	Variety
166	WH1402	SHORTENED SR26 TRANSLOCATION//2*WBLL1*2/KKTS/3/BECARD	Variety
167	BWL9022	C591/3*PBW343	Genetic stock
168	IC296440	LOCAL COLLECTION FROM GUJARAT	Genetic stock
169	SQ1	HIGHBURY/(TW269/9/3/4.CS)	Genetic stock
170	LGM245	LOCAL COLLECTION FROM GUJARAT	Landrace
171	LGM134-B	LOCAL COLLECTION FROM GUJARAT	Landrace

S. No.	Accession	Pedigree	Accession type
172	RIL116	DHARWAD DRY/DPW621-50	Genetic stock
173	NB29-6-2-12	LOCAL COLLECTION FROM GUJARAT	Landrace
174	C306	REGENT-1974/CH-23//2*C-591/3/P-19/C-281	Variety
175	IC321886	SK 41, LOCAL COLLECTION FROM BHOPAL, MP	Landrace
176	SB20	BH1146/HD2967	Genetic stock
177	BNI BORLAUG	Developed by integrating BNI trait from <i>Leymus racemosus</i> into Borlaug	Genetic stock
178	BH1146	Fronteira/Mentana//Ponta Grossa1	Variety

2.2. KASP genotyping of functional genes

A total of 40 KASP assays targeting SNPs in 35 functional genes related to adaptation, biotic and abiotic stress responses, grain quality, and yield were used for genotyping (**Supplementary Table S1**). Of these, 26 assays were previously reported by Rasheed et al. [56], while the remaining 14 were adopted from multiple studies [28, 31, 32, 35, 36, 42, 48]. These assays were designed based on diagnostic SNPs, most of which have been widely used for wheat germplasm characterization [36, 59, 60]. The nucleotide sequences of the KASP primers were previously reported in the studies above. However, we added a standard FAM tail (5'-GAAGGTGACCAAGTTCATGCT-3') and a HEX tail (5'-GAAGGTGCGAGTCAACGGATT-3') to facilitate fluorescence-based allele discrimination. The primers were synthesised by Barcode Biosciences Pvt. Ltd. (Bengaluru, India). The details of the SNP loci, along with the KASP primer sequences, are provided in Table 2. The KASP genotyping assays were performed in a 5.0 µL reaction volume containing 2.2 µL of DNA at a concentration of 40 ng/µL, 2.5 µL of 1x KASP V4.0 2X Master Mix (KBS-1016-017), 0.04 µL of Mg²⁺, 0.056 µL of primer mixture, and 0.204 µL of double-distilled water (ddH₂O). Thermal cycling conditions included an initial denaturation at 95°C for 15 minutes, a touchdown phase of 10 cycles at 95°C for 20 seconds and 65°C for 25 seconds (decreasing by 1°C per cycle), followed by an annealing phase of 30 additional cycles at 95°C for 10 seconds and 57°C for 60 seconds. KASP genotyping analysis was performed using the SNPliner™ workflow (LGC Genomics), which involved PCR amplification, fluorescent microplate reading, and cluster calling for genotype assignment. Fluorescence signals were recorded using a FLUOstar Omega SNP plate reader and imported into Kluster caller version 2.22.0.5 (LGC Hoddlesdon, UK) for allelic discrimination and genotyping analysis. Final genotype calls were exported as a CSV file.

Table 2
List of KASP-SNP primers used in the study.

Trait	Gene	Forward primer (5 – 3') (FAM)	Forward primer (5 – 3') (HEX)	Reverse primer (5'-3')
Plant height	<i>Rht-B1</i>	CCCATGGCCATCTCCAGCTG	CCCATGGCCATCTCCAGCTA	TCGGGTACAAGGTGCGGG
	<i>Rht-D1</i>	CATGGCCATCTCGAGCTGCTC	CATGGCCATCTCGAGCTGCTA	CGGGTACAAGGTGCGCGC
Vernalization	<i>Vrn-A1</i>	AGAGTTTTCCAAAAAGATAGATCAATGTAAAT	GAGTTTTCCAAAAAGATAGATCAATGTAAAC	GTTAGTAGTGATGGTCCA/
	<i>Vrn-B1</i>	GGCAGCTAATGTGGGGTAGTCA	GGCAGCTAATGTGGGGTAGTCT	CACAGGCTTTCTATCATI
	<i>Vrn-D1</i>	ATCATTCTGAATTGCTAGCTCCGG	ATCATTCTGAATTGCTAGCTCCGC	GCCTGAACGCCTAGCCTG
Photoperiod sensitivity	<i>Ppd-D1</i>	CAAGGAAGTATGAGCAGCGGTT	AAGAGGAAACATGTTGGGGTCC	GCCTCCCACTACACTGGG
Pre-harvest sprouting	<i>PHS1_646</i>	GGTGGAACAGATGCAACTAAAGG	GGTGGAACAGATGCAACTAAAGA	GTGAGTGTTATATGAAACI
	<i>PHS1_666</i>	GTGAGTGTTATATGAAACTAATGATCCATTT	GTGAGTGTTATATGAAACTAATGATCCATTA	ACCGGGTGGAAACAGATGC
	<i>Sdr-B1</i>	TCAGCAGACTTCGACTCGCA	TCAGCAGACTTCGACTCGCG	AGGATCTCGTTGGCCTTG/
	<i>MFT_1617R</i>	CGTGCATGCAGCCTACGCATACGTA	CGTGCATGCAGCCTACGCATACGTG	TGATCCATGCACGCATCA(
	<i>MFT_721J</i>	GTCCTTCCAGAGGCCATGTCTG	GTCCTTCCAGAGGCCATGTGC	ACAGGTCAAGCGACACGT
	<i>VP-1B</i>	CGCTCTTATATTAGTTTACGGAGGGAG	CTCATTTTGATAGCTCTAGCTAA	CTTCTTCCGAAGTGATCA/
	<i>PH2L_SNP</i>	AGCCCGGAATCCTTCGCA	AGCCCGGAATCCTTCGCG	CGTCAACCTTCTTCACCC'
	<i>PH4L_SNP</i>	TGACGATTATTCCAGTGACAGATA	TGACGATTATTCCAGTGACAGATG	GCCAAGGACACGATGGTT
	<i>PH5L_SNP</i>	GCGTACAGCCACACTTGT	GCGTACAGCCACACTTGC	GGTTAAGATGGTGTATCA(
	<i>PH6L_SNP</i>	GTGTGGCTCAATGGGATGACA	GTGTGGCTCAATGGGATGACG	AACCAACCGGCTGCACCA
	<i>Dreb-B1</i>	CCTGCGCACTTTCTTCTTCTGT	CTGCGCACTTTCTTCTTCTGG	TTTCACCTTGATATGGA
	<i>Lr14a</i>	CTACACTAGTACTACTTTGAGACAATTTTTT	ACACTAGTACTACTTTGAGACAATTTTAA	AACAAACTCCAGTGTAAG/
Leaf rust resistance	<i>Lr34</i>	GGTATGCCATTTAACATAATCATGAA	GGTATGCCATTTAACATAATCATGAT	TACTATATGGGAGCATTAT
	<i>Lr68</i>	CGTGTCTTGGACCTGAGCAAT	CGTGTCTTGGACCTGAGCAAC	TGACCTGAGTCCCGTCAA
	<i>Yr15_R8</i>	CAGATCCCCGGTTCTCTCAAG	CAGATCCCCGGTTCTCTCAAA	CCCCCAAATGATCGAGAA'
Stripe rust resistance	<i>Yr15_R5</i>	AGTCAACTTGGATTACACTGAAGTT	AGTCAACTTGGATTACACTGAAGTC	AGATATCACACTGAACAT/
	<i>Fhb1</i>	CACATGCATTTGCAAGGTTGTTATCC	CACATGCATTTGCAAGGTTGTTATCG	CAAAGCAGCCTTAGGTCA
Fusarium head blight	<i>PFT_KASP</i>	AGCTGGGAATGAAGGAACCTGC	AGCTGGGAATGAAGGAACCTGT	GTGATTATCTCTCCCATCT
	<i>1BL/1RS</i>	GGAGCAGGTCCAGATCGCG	CGGAGCAGGTCCAGATCGCA	GAAGCTCCGGTAGATGGA
Quality	<i>Glu-A1</i>	AAGTGTAACCTTCTCCGCAACG	ACCTAAGTGTAACCTTCTCCGCAACA	CGAAGAAGCTTGGCCTGG
	<i>Glu-D1</i>	ATAGTATGAAACCTGCTGCGGAG	ATAGTATGAAACCTGCTGCGGAC	TACTAAAAAGGTATTACCC
Quality	<i>Pina-D1</i>	AAGTCCCAACAACCTTCGCTA	TTGTCTAGTACCCCGCTCTG	ATGAAGGCCCTCTTCTCT/
	<i>Pinb-D1</i>	CTCATGCTCACAGCCGCC	CCTCATGCTCACAGCCGCT	GTCACCTGGCCACAAAA
	<i>Pinb2-B2</i>	GCACCTAGCAATAAATAAACGGGAG	AGAAAAAGCCATTAAATAAACGGGAC	TGTTTTGGTGGTGGTGAA/
Sucrose synthase	<i>Sus1-7A-1185</i>	GATTTGATCCATGCCCTCTC	GATTTGATCCATGCCCTCTT	CTGTCTTCAACATCATTC
	<i>Sus1-7A-1599</i>	CGGTATCAGGGTATCCCAACACATTGGA	CGGTATCAGGGTATCCCAACACTTTGTT	ATCCTGTCTCCCATGGA1
	<i>Sus1-7B</i>	CAATTGCTTATGTTCTGTTGTATGG	CAATTGCTTATGTTCTGTTGTACAT	ATGTTTATGCTTGAATGG/
	<i>Sus2-2A</i>	AGCAATGGGGGAGACTGCTGGAGG	AGCAATGGGGGAGACTGCTGGAGA	GCTGTGGATGCGGCTCAG
Cell wall invertase	<i>Cwi-4A</i>	TTTATTTAAAAATTTGATGAACTTTTCATAAAC	TTTATTTAAAAATTTGATGAACTTTTCACAAAT	CATCGAATTGAAGAAAAA'
Grain width	<i>GW2-6A</i>	GGTCGATGAGATCCCGTACAA	GGTCGATGAGATCCCGTACAG	CGAGCTAGGGTTTGTGTC
	<i>GW2-6B</i>	GTTGGTGTCAATTTGTAAAGCCCA	GTTGGTGTCAATTTGTAAAGCCCC	AGGCTTGTCAAACGTGG

Trait	Gene	Forward primer (5 – 3') (FAM)	Forward primer (5 – 3') (HEX)	Reverse primer (5'-3')
Grain weight and length	<i>GS-D1</i>	GCCAAGAAATGTCGCTCTCAG	GCCAAGAAATGTCGCTCTCAT	CAAGAATTTTGGGACGGA
Grain weight	<i>TaSBEIII-A</i>	ACTGCTCTCCGCCGGTCCGGCTGCC	ACTGCTCTCCGCCGGTCCGGCTGCT	CCACAACCTCCTGCCAGG
Spikelet number per spike	<i>MOC-7A</i>	CGGAAAGCTTAGAAATGCACGG	CGGAAAGCTTAGAAATGCACGT	CACCTCTAATCCAATGCGA

Note: Standard FAM (5 GAAGGTGACCAAGTTCATGCT-3) and HEX (5-GAAGGTGCGAGTCAACGGATT-3) tails with the targeted SNP not included with the primers.

2.3. Functional allele matrix construction

Genotypic data from 40 KASP loci were curated to create a binary functional allele matrix, where the presence of the favourable allele was scored as “1,” and the absence or alternate alleles were scored as “0.” The curated matrix served as input for analyses of diversity, structure, and zone suitability.

2.4. Population structure and phylogenetic analysis

A Neighbour-Joining (NJ) tree was constructed using DARwin v6.0 [61] to clarify the phylogenetic relationships among accessions. Principal Coordinate Analysis (PCoA) was performed using the adegenet v2.0.1 package in R [62], and the first three principal coordinates, which explained most of the genetic variation, were used to visualize population structure. Bayesian inference of population structure was performed using STRUCTURE version 2.3.4 [63] under an admixture model with correlated allele frequencies. Each STRUCTURE run included a burn-in period of 100,000 iterations followed by 500,000 Markov Chain Monte Carlo (MCMC) iterations, with 10 independent replicates for each assumed number of populations (K) from 1 to 10. The optimal number of genetic clusters was determined using the ΔK method of Evanno et al. [64] and STRUCTURE HARVESTER. Allelic variation parameters, such as the number of alleles per locus and allele frequencies, were calculated, and genetic diversity indices, including expected heterozygosity (H_e), observed heterozygosity (H_o), gene diversity, and polymorphic information content (PIC), were estimated using PowerMarker version 3.25 [65]. Differences in genetic diversity between two predefined genotype groups were tested using Student’s t-test, with $P < 0.05$ considered significant.

2.5. Breeding zone definitions and trait prioritization

The five agro-climatic wheat production zones of India are the North Western Plain Zone (NWPZ), the North Eastern Plain Zone (NEPZ), the Central Zone (CZ), the Peninsular Zone (PZ), and the Northern Hills Zone (NHZ) (Directorate of Wheat Development, 2016). Zone-specific breeding priorities were compiled from technical bulletins and guidelines published by the ICAR-Indian Institute of Wheat and Barley Research, Karnal – 132 001, Haryana (India) [66, 67], as well as consultations with experts.

2.6. Construction of the rule-based trait–zone matrix

Zone-specific breeding priorities were formalized using a rule-based trait–zone matrix to enable trait-weighted genotype selection. Each trait was assigned a zone-specific weight proportional to its relative breeding priority (**Supplementary Table S2**). These weights were subsequently mapped to the functional loci associated with each trait analyzed in the study. For traits governed by multiple loci, the zone-level trait weight was distributed equally among all contributing loci. This approach was adopted because quantitative estimates of locus-specific effect sizes are not consistently available across Indian wheat production environments, and equal distribution reduces bias resulting from incomplete knowledge of allelic effects.

2.7. Zone-specific genotype-suitability scores calculated using a weighted, allele-based model

For each genotype g , locus l , trait t , and production zone z , allele states were encoded as $a_{g,l} \in \{0,1\}$ indicating the absence or presence of the favourable allele. Trait level weight ($W_{z,t}$) were obtained from zone-specific breeding priorities, with expert consultation. When multiple loci controlled a trait (L_t), its weight ($W_{z,t}$) was split equally across the corresponding loci, such that each locus received

$$w_{z,l} = \frac{W_{z,t}}{|L_t|}, \quad l \in L_t$$

If locus l is not assigned to any trait, set $w_{z,l} = 0$

The raw score for genotype g in zone z was computed as the weighted sum of favourable alleles:

$R_{g,z} = \sum_{l \in L} a_{g,l} w_{z,l}$ Suitability score was then normalized to a 0-100 scale using the total zone-specific weight:

$$S_{g,z} = \frac{R_{g,z}}{\sum_{l \in L} w_{z,l}} \times 100$$

Genotypes were ranked within each zone based on their $S_{g,z}$ -scores. For each zone, the five accessions with the highest scores were assigned to that zone. To maintain a broad genetic base, each genotype was assigned to only one zone. In cases where no genotype met the maximum-score criterion for a zone—due to prior assignment to another zone—the genotype with the next highest score was selected to ensure full zone representation.

3. Results

3.1. Genotyping and population structure

The genotypic analysis of 178 bread wheat genotypes using 40 functional KASP assays revealed polymorphism at all loci, each showing high allelic resolution and distinct clustering patterns (Fig. 1). The loci were functionally classified into four categories, including grain yield (10), grain quality (5), adaptation (16), and stress resistance (9), based on prior validations.

Neighbour-joining (NJ) clustering (Fig. 2a) and principal component analysis (PCA) (Fig. 2b) consistently partitioned the panel into two major genetic groups. STRUCTURE analysis further confirmed $K=2$ as the optimal number of subpopulations according to the ΔK method (Fig. 2c). The first cluster (Group A) predominantly comprised semi-dwarf, high-yielding Mexican-derived cultivars that are responsive to intensive management and exhibit broad biotic stress resistance. The second cluster (Group B) mainly included taller, traditional Indian and indigenous cultivars characterised by lower input responsiveness but superior tolerance to abiotic stresses such as heat and drought, albeit with relatively higher susceptibility to biotic stresses. These results delineate two complementary gene pools shaped by contrasting selection histories and breeding objectives.

3.2. Genetic diversity in the population

Genetic diversity parameters were estimated across the 40 functional loci to quantify polymorphism within the panel. Substantial allelic variation was detected, confirming a broad genetic base and considerable functional divergence among genotypes. The highest gene diversity ($H_e = 0.498$) was observed at *Rht-B1*, *PH4L*, and *PHS1_666*, all of which are associated with adaptation, indicating strong allelic heterogeneity and possible balancing selection that maintains adaptive potential. Conversely, *Yr15_R5* (biotic stress), *PFT_KASP* corresponding to *Fhb1* (biotic stress), and *GW2-6B* (yield) exhibited the lowest diversity ($H_e = 0.022$), suggesting fixation or near-fixation of favourable alleles due to directional selection or introgression in elite breeding lines.

Across functional categories, adaptation-related loci showed the highest mean diversity ($H_e = 0.420$), followed by stress-resistance ($H_e = 0.325$), quality ($H_e = 0.275$), and yield-related genes ($H_e = 0.263$). This gradient indicates stronger selection pressure on yield and quality traits compared with the broader variability retained in adaptive loci. Minor-allele frequency analysis (allele frequency < 0.5) revealed that landraces (mean = 9.35) harboured more rare alleles than improved varieties (8.75) or genetic stocks (7.76), identifying them as valuable reservoirs for pre-breeding and diversity enhancement.

3.3. Allelic frequency patterns across genetic groups

Comparative analysis of allelic frequencies between the two STRUCTURE-defined subpopulations revealed distinct yet interconnected selection patterns across adaptive, stress-tolerance, quality, and yield-related loci (Fig. 3).

3.3.1. Adaptation-related loci

Both subpopulations exhibited broadly similar allelic profiles, with overlapping interquartile ranges and only a slightly higher median frequency of favourable alleles in Group A. Nevertheless, locus-level differentiation was evident (Fig. 4a). A significant divergence occurred at *Rht-D1*, *Vrn-B1*, *PH-4L*, *PH-6L*, *PHS1_646*, *PHS1_666*, *Sdr-B1*, *Vp-1B*, and *MFT-A1*, whereas loci such as *Rht-B1*, *Vrn-A1*, *Vrn-D1*, *Ppd-D1*, *PH-2L*, and *PH-5L* remained largely undifferentiated. Group A was enriched for favourable alleles at *PH-4L*, *PH-6L*, *PHS1_646*, *PHS1_666*, and *Sdr-B1*, while Group B showed enrichment at *Rht-D1*, *Vrn-B1*, *Vp-1B*, and *MFT-A1*. These patterns suggest a polygenic, heterogeneous basis of adaptation, shaped by localized selection and differential introgression.

3.3.2. Stress-resistance loci

For biotic and abiotic stress-associated loci, Group A displayed slightly higher median allele frequencies, consistent with deliberate introgression of multiple resistance sources during breeding (Fig. 4b). Locus-specific patterns revealed enrichment of favourable alleles at *Dreb-B1*, *Yr15_R5*, and *Yr15_R8* in Group A, whereas *Lr68*, *Fhb1* (*PFT_KASP*), and *1BL.1RS* were more frequent in Group B. Although *Fhb1* exhibited low overall diversity, its enrichment within Group B reflects differential fixation of resistance alleles in distinct breeding pools. Meanwhile, *Lr14a* and *Lr34* exhibited nearly identical allele frequencies across both groups, indicating the widespread maintenance of these core resistance loci.

3.3.3. Grain-quality loci

Grain-quality loci exhibited overall homogeneity, characterised by nearly identical medians and overlapping distributions across the two groups (Fig. 4c). This pattern reflects successful introgression and fixation of major quality alleles throughout the population. However, minor locus-specific deviations persisted: *Glu-D1* was more frequent in Group A, whereas *Pinb2-B2* was enriched in Group B. The *Glu-A1*, *Pina-D1*, and *Pinb-D1* variants exhibited balanced representation, reflecting minor population-specific divergence within a largely homogenised genetic background.

3.3.4. Yield-related loci

Allele frequency distributions for yield-related loci were also largely balanced between the two subpopulations (Fig. 4d). Group A exhibited slightly higher median frequencies but substantial overlap with Group B, confirming the widespread dissemination of productivity-enhancing alleles through recurrent selection. Locus-level differences included enrichment of *Gw2-6A* in Group A and *TaSBEIIIA*, *MOC-7A*, *Sus1-7A-1599*, and *Sus1-7B* in Group B, whereas *GS-D1*, *GW2-6B*, *Sus1-6A-1185*, *Sus2-2A*, and *Gwi-4A* showed near parity. These patterns demonstrate a genomic architecture shaped by both convergence through modern breeding and subtle locus-specific divergence that preserves adaptive potential.

3.4. Contribution of founder parents to favourable allele distribution

Seven key founder parents (Kalyansona, KRL1-4, Sonalika, WH542, WL711, YR10, and YR15) were evaluated across the diversity panel to trace the origins and accumulation of favourable alleles (Fig. 5). Distinct yet complementary patterns emerged between Groups A and B, reflecting differences in founder usage and selection intensity.

3.4.1. Adaptation-related loci

In Group A, favourable alleles were evenly distributed across multiple allele classes, with peaks at six-allele classes where both founder and diverse parental lines contributed comparably (Fig. 5a). This balanced distribution indicates that recombination among founders and diverse lines facilitated the gradual enrichment of adaptive alleles, promoting a broad-based polygenic adaptation. In contrast, Group B (Fig. 5b) showed a more skewed pattern, dominated by founder parents in the five-allele class and diverse lines in the six-allele class, suggesting a more directional selection history aimed at defined agro-ecological adaptation.

3.4.2. Stress-related loci

For stress-resistance loci, Group A exhibited a founder-dominated pattern with a right-skewed distribution peaking at four alleles (Fig. 5c). This indicates that resistance genes were primarily introduced through deliberate introgression from specific donors. Group B (Fig. 5d), however, showed a broader, higher distribution peaking at five alleles, with founders and diverse genotypes contributing equally to the modal class. The presence of genotypes carrying six or seven favourable alleles highlights the cumulative pyramiding of stress-resistance loci through recombination and selection, reflecting a more integrated breeding pipeline.

3.4.3. Grain-quality loci

Grain-quality alleles showed minor but distinct differentiation (Fig. 5e). In Group A, genotypes were concentrated in the lowest favourable-allele class, with founder parents accounting for about half of this group, indicating recurrent use of elite lines to fix key quality alleles. Group B (Fig. 5f) exhibited a more balanced distribution across allele classes, with diverse genotypes occupying the highest-allele class, highlighting rare, founder-independent combinations that may broaden the genetic base for end-use quality traits.

3.4.4. Yield-related loci

Yield-associated alleles were widely distributed, with modal classes of five to seven favourable alleles in both groups (Fig. 5g). Group A peaked at six alleles, represented mainly by founder-derived genotypes, while Group B (Fig. 5h) displayed a flatter distribution with greater contribution from diverse lines in the modal class. The rare eight-allele carriers occurred exclusively among diverse genotypes, suggesting untapped yield potential within non-founder germplasm.

Across all analyses, the genomic pattern of Indian bread wheat shows two complementary routes of breeding improvement. Group A represents a founder-centric approach that emphasizes the fixation of major alleles for adaptation, stress tolerance, and yield through targeted selection. Group B represents a recombination-driven approach that incorporates alleles from improved cultivars, landraces, and introgression lines to increase genetic diversity. Founder-based selection has stabilized key agronomic traits, whereas recombination among diverse parental backgrounds has generated novel and stacked allele combinations, particularly for stress resilience and quality.

3.5. Zone-wise identification of elite wheat genotypes using weighted allelic suitability scores

Using zone-specific suitability scores derived from weighted favourable-allele profiles across 40 KASP assays targeting SNPs in 35 functional genes related to adaptation, biotic and abiotic stress responses, grain quality, and yield, zone-wise rankings were created for all 178 genotypes. Finally, five genotypes were selected for each of the six Indian wheat production zones as the most suitable candidates for targeted breeding based on the maximum-score criterion for each zone. Each genotype was assigned to only one zone to maintain genetic diversity across the national breeding pipeline. The top five highest-scoring, non-redundant genotypes for each of the six Indian wheat production zones are summarized in Table 3.

Table 3

Zone-wise prioritization scores and top five wheat genotypes based on zone-specific relative breeding priorities of evaluated traits.

S. No.	Zone	Criteria (ordered according to the zone-wise relative breeding priority of the evaluated traits)	Genotype	Score
1	NWPZ	Stripe rust resistance; Leaf rust resistance; Grain yield; Reduced height; Photoperiod sensitivity; Grain quality; Vernalization; Pre-harvest sprouting; Drought tolerance; Fusarium head blight resistance	PBW723	75.8
2			HD2967	73.6
3			VL798	72.9
4			D127	72.1
5			DBW371	64.9
6	PZ	Drought tolerance; Photoperiod sensitivity; Grain yield; Stripe rust resistance; Leaf rust resistance; Grain quality; Pre-harvest sprouting; Reduced height; Vernalization; Fusarium head blight	DHARWADDRY	83.0
7			FLW13	77.8
8			PNBW596	76.0
9			W485	75.8
10			MACS6222	75.2
11	NHZ	Vernalization; Stripe rust resistance; Photoperiod sensitivity; Leaf rust resistance; Grain yield; Reduced height; Pre-harvest sprouting; Fusarium head blight resistance; Grain quality; Drought tolerance	IC321998	75.0
12			IC322007A	73.3
13			NP846	73.1
14			IC322001	72.9
15			HPW349	68.3
16	NEPZ	Drought tolerance; Photoperiod sensitivity; Pre-harvest sprouting; Leaf rust; Stripe rust; Grain quality; Grain yield; Reduced height; Vernalization; Fusarium head blight resistance	HD2888	72.2
17			FLW5	70.8
18			MBL5	66.5
19			HD3118	66.3
20			NW1014	66.2
21	CZ	Drought tolerance; Photoperiod sensitivity; Grain yield; Grain quality; Leaf rust resistance; Stripe rust resistance; Pre-harvest sprouting; Reduced height; Vernalisation; Fusarium head blight resistance	NARMADA4	72.3
22			HINDI62	72.2
23			HI1500	72.0
24			NB29-13-5-9	68.6
25			KBRL10	68.3

4. Discussion

A comprehensive understanding of functional gene diversity is crucial for the genetic improvement in bread wheat (*Triticum aestivum* L.). Recent advances in high-throughput genotyping, molecular breeding, and functional genomics have equipped the breeders to dissect, monitor, and manipulate key loci governing adaptation, resistance traits, end-use quality, and productivity with high precision [36]. The genetic architecture of modern wheat cultivars reflects the cumulative integration of landraces, founder parents, and introduced germplasm, each contributing distinct allelic combinations that collectively reinforce agronomic performance and stress adaptability [68]. Among these, landraces remain invaluable reservoirs of adaptive allelic variation, having co-evolved under diverse environmental and management conditions across centuries. Their long-term natural and farmer-mediated selection has retained broad-spectrum alleles conferring stress resilience, phenological plasticity, and local adaptation, traits often diminished through intensive selection in elite breeding pools [69, 70]. Conversely, introduced germplasm, particularly materials carrying key alleles such as *Rht-B1b* and *Rht-D1b* (semi-dwarfing genes) [71], the 1BL.1RS rye translocation (multi-trait complex conferring disease resistance and yield advantage) [72], and yield-enhancing loci such as *GW2-6B* and *Sus2-2B*, has expanded the allelic range of modern cultivars and reinforced their genetic foundation for yield stability and adaptation [36]. In self-pollinated polyploid crops like wheat, where genome complexity and low recombination rates pose major challenges, high-throughput functional genotyping platforms have transformed the precision and scale of genetic dissection. Among these, Kompetitive Allele-Specific PCR (KASP) marker assays targeting validated functional gene variants represent one of the most robust and cost-effective tools for allelic discrimination [56]. By enabling accurate, biallelic genotyping across large populations, KASP assays have accelerated the identification and introgression of favourable alleles, enhanced marker-assisted backcrossing efficiency, and facilitated the pyramiding of complementary genes for complex traits. The integration of such marker systems into breeding pipelines now reinforces a new era of data-driven selection, wherein allelic diversity can be systematically harnessed to maximise both genetic gain and adaptive potential in wheat improvement programs [73].

In the present study, genotyping of the Indian bread wheat panel using KASP markers revealed distinct allelic variation and a well-defined population structure, partitioning the germplasm into two primary genetic clusters that reflect distinct breeding lineages and agro-ecological adaptation patterns. Such

structuring highlights the dual influence of targeted selection and localised adaptation in shaping the current allelic composition of Indian wheat [74]. Globally, the evolution of modern breeding, characterized by recurrent use of elite founder lines and strong directional selection for yield, height reduction, and quality traits, has led to a significant erosion of overall genetic diversity within improved cultivars compared to their ancestral landraces [4]. Nonetheless, landraces remain a vital repository of allelic richness, particularly for genes associated with environmental adaptability, stress tolerance, and phenological flexibility [28]. This persistent genetic heterogeneity not only preserves valuable alleles lost during domestication and intensive breeding but also provides a critical foundation for pre-breeding programs aimed at reintroducing adaptive diversity into elite pools [75, 76]. Consequently, strategic exploitation of landrace diversity through genomic-assisted introgression, multi-parental populations, and targeted allele mining represents an essential pathway for broadening the functional gene base and sustaining long-term genetic gains in global wheat improvement [77, 78].

The population structure observed in our diverse wheat panel reflects the convergent influence of founder-derived fixation and recombination-driven diversification, two complementary forces that have jointly shaped the genomic architecture of modern cultivars. The founder-derived group (Group A) exhibits a high frequency of fixed major alleles associated with key agronomic and adaptive traits, consistent with the recurrent use of a limited set of elite parental lines in breeding programs. This pattern typifies the genetic consequences of selective breeding bottlenecks, wherein desirable alleles for yield stability, height regulation, and quality are preferentially retained through intensive selection and backcrossing. In contrast, the recombination-driven group (Group B) represents a dynamic genetic mosaic, characterized by novel and stacked allelic combinations arising from hybridization among improved cultivars, landraces, and wild relatives. Such allele reshuffling enhances the potential for adaptive innovation and broadens the functional diversity available for improving complex traits [79]. Evidence from MAGIC populations shows that both founder effects and recombination contribute to wheat improvement. These populations increase recombination and allow precise mapping of complex traits across founder lines [80]. Thus, trait variation and new allele combinations in wheat arise from the dual action of founder haplotypes and recombination, ensuring stable productivity and adaptability in modern germplasm.

The present study provides clear evidence of differential selection pressures across functional gene classes in bread wheat, reflecting the combined influence of breeding priorities and agro-ecological adaptation. A distinct gradient in allelic diversity, highest at adaptation-related loci and lowest at loci associated with yield and quality, demonstrates that selection intensity varies among trait categories. This pattern of hierarchical diversity reflects the historical course of wheat improvement, where selection for agronomic uniformity and productivity progressively reduced allelic variation at major-effect loci while maintaining greater diversity at adaptive genes. Strong, directional selection for semi-dwarf stature mediated by the *Rht-B1b* and *Rht-D1b* alleles, photoperiod and vernalization responsiveness determined by *Ppd* and *Vrn* loci, and key end-use quality determinants such as *GS-D1* and *Pinb-D1* have led to near-complete fixation of favourable alleles within elite breeding germplasm [21, 22, 81]. These loci have undergone strong artificial selection due to their major phenotypic effects and predictable contributions to yield potential, grain quality, and agronomic stability, resulting in reduced allelic polymorphism and constrained genetic plasticity in advanced cultivars [36, 82]. Conversely, adaptation-related genes, which govern physiological plasticity and environmental responsiveness, exhibited markedly higher allelic heterogeneity. Their diversity reflects the cumulative effects of region-specific selection, introgression from locally adapted landraces, and recurrent recombination events introduced during breeding cycles targeting diverse Indian agro-ecologies [74]. Such heterogeneity sustains adaptive potential and helps maintain cryptic variation essential for resilience under variable climatic regimes. Comparable diversity gradients across global wheat populations confirm that diversity erosion is locus-specific and governed by the direction, strength, and timing of selection [75, 83]. These findings emphasize that modern wheat breeding, while enhancing productivity and uniformity, must strategically balance the fixation of elite alleles with the conservation and reintegration of adaptive variation to sustain long-term genetic and agronomic resilience.

As functional allelic diversity becomes increasingly well-characterized, its application in breeding requires operational strategies that effectively transform allele-level information into actionable selection decisions. To address this, zone-targeted deployment of alleles based on trait-diagnostic markers establishes a direct link between molecular diversity characterization and environment-specific breeding strategies. Currently, breeders have access to an expanding catalog of validated functional polymorphisms for traits such as height, phenology, disease resistance, drought adaptation, and grain quality. Utilising KASP-based assays, breeders can track these favourable alleles with high accuracy and cost-efficiency across diverse germplasm sets. Building on these resources, this study advances trait-diagnostic markers from basic diversity assessment tools to a quantitative decision-support system for zone-wise genotype prioritization. Specifically, favourable-allele states at 40 SNPs across 35 functional genes were encoded as a binary matrix, with each locus weighted by zone-specific breeding priorities. For example, stripe and leaf rust resistance received priority in the North Western Plains Zone (NWPZ), while drought and phenology traits were emphasized in the Peninsular (PZ) and Central Zones (CZ). Similarly, vernalization was prioritized in the Northern Hill Zone (NHZ). Zone-wise suitability scores were calculated for 178 genotypes and normalized to a 0-100 scale. Five elite candidates per zone were selected using a maximum-score rule, with each genotype limited to a single zone to minimize redundancy and preserve allelic diversity in national breeding programs. This marker-based suitability index differs from conventional marker-assisted selection, which typically tracks a limited number of major genes in isolation, and from genomic selection, which optimizes aggregate breeding value without explicit zone-wise trait priorities. The present approach integrates trait-diagnostic loci with established phenotypic effects, expert-informed weightings for each zone, and a specific rule for genotype allocation. Previous applications of diagnostic markers have focused on targeted introgression, pyramiding of major genes, or diversity studies. However, there is no evidence in the wheat literature of an operational framework that converts functional alleles into weighted, zone-specific suitability scores for allocating genotypes according to official breeding priorities. To the best of current knowledge, this study presents a novel operational pipeline that converts functional alleles into weighted, zone-aware suitability scores for genotype deployment. This approach is generalizable; as additional functional loci are validated, the index can be updated, enabling iterative re-prioritization of breeding materials and establishing a transparent link between functional genomics and environment-specific improvement programs.

The enrichment of rare alleles and unique functional variants within landraces and germplasm collections constitutes an essential reservoir of novel genetic variation for resilience breeding. These alleles, often lost or underrepresented in modern cultivars due to historical selection bottlenecks, possess adaptive potential for traits such as biotic and abiotic stress tolerance and nutrient-use efficiency that are increasingly important under changing climatic and production conditions. Khalid et al. [59] reported that landrace-derived germplasm harbours a higher frequency of rare favourable alleles than elite breeding

lines, particularly at loci associated with stress adaptation and physiological plasticity. Such allelic richness offers a foundation for expanding the functional gene space of modern wheat through targeted introgression and allele mining. Historical examples, such as the rye-derived 1BL.1RS translocation and wild-relative introgressions conferring stress tolerance [84], demonstrate the impact of genetic base broadening on wheat resilience and productivity. Efficient utilization of these alleles requires robust genotyping and analytical procedures such as multiplexed KASP marker panels, high-density SNP arrays, and haplotype-based analyses to identify favourable allelic combinations while minimizing linkage drag. Integrating these genomic tools within structured pre-breeding pipelines enables systematic transfer of novel alleles into elite backgrounds, followed by selection schemes that eliminate undesirable linkages and restore high agronomic performance. This approach ensures that genetic diversity derived from landraces and exotic germplasm is effectively translated into functional resilience and yield stability without compromising adaptation to modern production systems.

The genetic architecture of yield and quality traits in wheat reflects both the achievements and continuing challenges of modern breeding. Key loci governing end-use quality, such as *Glu-D1* and *Pinb2-B2*, have undergone intense positive selection, resulting in near-fixation of favourable alleles in many Indian elite germplasm pools [81, 85]. Nonetheless, considerable residual variation persists, reflecting incomplete fixation and the influence of local adaptation. This residual diversity presents opportunities for fine-scale mapping of environmentally responsive alleles and for optimising allelic combinations tailored to specific production environments. The prevailing challenge is thus not merely the universal fixation of a single “best” allele but the strategic deployment of alleles in accordance with target agro-ecological contexts. Such precision breeding is increasingly feasible through the integration of high-throughput functional genotyping platforms, large-scale multi-environment phenotyping networks, and genomic prediction models that explicitly incorporate genotype × environment (G×E) interactions to enhance predictive accuracy and selection efficiency [86].

Overall, the present study emphasizes the strategic integration of founder-derived stability with diversity-driven innovation to achieve durable genetic improvement in bread wheat. Founder alleles provide a foundation for yield stability, adaptation, and quality, while rare alleles from landraces and exotic germplasm contribute novel variants essential for resilience, nutrition, and quality enhancement. Integrating these complementary allele pools through multi-environment phenotyping, functional validation, and precise introgression enables balanced genetic advancement. High-throughput functional genotyping platforms such as KASP, when combined with genomic selection and multi-parent populations, offer an efficient strategy for pyramiding validated loci and optimising polygenic background merit. Functional gene diversity thus represents a central determinant of sustained genetic gain, enabling the assembly of superior allele combinations that enhance both productivity and environmental adaptability.

5. Conclusion

Functional gene diversity is recognized as a key factor in adaptive capacity and sustained genetic gain in bread wheat. The balance between founder-derived stability and recombination-driven innovation forms the genetic foundation of modern wheat breeding. Founder alleles confer agronomic stability and quality, whereas landraces and exotic derivatives introduce rare adaptive variants that are vital for stress resilience and environmental adaptability. The gradient of allelic diversity across functional categories reflects the impact of historical selection and breeding goals. High-throughput KASP-based functional genotyping allows precise identification of allelic states and supports the development of reliable trait-diagnostic markers for region-specific allele deployment. Matching adaptive alleles with regional agroecological conditions offers a biologically informed strategy to optimise genotype–environment interactions. Future breeding progress will rely on integrating functional marker-assisted selection with genomic selection to capture both major-effect loci and polygenic background variation. Incorporating landrace diversity, targeted introgression of adaptive alleles, and extensive marker validation across diverse environments are essential for developing wheat cultivars with high productivity, improved nutritional quality, and resilience to climatic and biotic stresses, thereby ensuring durable genetic improvement and long-term food security.

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with all ethical guidelines and principles and was approved by the competent authority of the Institute. All participants provided informed consent before participating in the study.

Consent for publication

All authors have agreed to the publication of this manuscript. The authors declare that they have no competing interests.

Competing interests

The authors declare no competing interests.

Funding

This study was conducted at the ICAR–Indian Institute of Wheat and Barley Research (IIWBR), Karnal–132 001, Haryana, India. Financial support for this work was provided by the CABin project of the ICAR–Indian Agricultural Statistics Research Institute (IASRI), New Delhi–110 012, India.

Author Contribution

B.K.S designed the experiments and wrote the manuscript. R.S. and N.R. performed the experiments. S.S., R.S., R.S.S., S.J., and M.A.I. carried out the data analyses. O.P.A. and R.T. conceived and coordinated the project. All authors reviewed and approved the final version of the manuscript.

Acknowledgement

The authors gratefully acknowledge the institutional support provided by the ICAR–Indian Institute of Wheat and Barley Research (IIWBR), Karnal, Haryana, India.

Data Availability

No new whole-genome or transcriptome sequencing data were generated in this study. All functional KASP markers analyzed were derived from previously published genes. Detailed information for each functional locus, including marker identity and associated traits, is provided in ****Supplementary Table S1****. Zone-wise trait weightings (%) used for functional allele–based genotype scoring across the five Indian wheat production zones are also provided in ****Supplementary Table S2****. All other data supporting the conclusions of this study are included within the article and its tables and figures.

References

1. Sears ER. Homoeologous chromosomes in *Triticum aestivum*. Genetics. 1952;37:624.
2. Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Consortium IWGS, Jakobsen KS, Wulff BBH, Steuernagel B, Mayer KFX, Olsen O-A. Ancient hybridizations among the ancestral genomes of bread wheat. Science. 2014;345:288–91. <https://doi.org/10.1126/science.1250092>.
3. Li LF, Zhang ZB, Wang ZH, Li N, Sha Y, Wang XF, Ding N, Li Y, Zhao J, Wu Y, Gong L, Mafessoni F, Levy AA, Liu B. Genome sequences of five Sitopsis species of *Aegilops* and the origin of polyploid wheat B subgenome. Mol Plant. 2022;15:488–503. <https://doi.org/10.1016/j.molp.2021.12.019>.
4. Feldman M. The origin of cultivated wheat. In: Bonjean A, Angus W, editors. The World Wheat Book. Paris: Lavoisier Tech. & Doc.; 2001.
5. USDA. (2024) <https://fas.usda.gov/data/production/commodity/0410000>
6. Tomar SMS, Vinod, Singh B. Distant hybridization in wheat. New Delhi, India: IARI; 2004.
7. Pal BP. (1966) Wheat. Indian Council of Agricultural Research (ICAR), New Delhi, India.
8. Poehlman JM, Borthakur D. Breeding Asian Field Crops. Oxford & IBH Publishing Co; 1959.
9. Swaminathan MS. Genesis and growth of the yield revolution in wheat in India: Lessons for shaping our agricultural destiny. Agric Res. 2013;2:183–8. <https://doi.org/10.1007/s40003-013-0069-3>.
10. Tadesse W, Manes Y, Singh R, Payne T, Braun H-J. Adaptation and performance of CIMMYT spring wheat genotypes targeted to high rainfall areas of the world. Crop Sci. 2010;50:2240. <https://doi.org/10.2135/cropsci2010.02.0102>.
11. Yadav R, Gaikwad K, Bhattacharyya R. Breeding wheat for yield maximization under conservation agriculture. Indian J Genet Plant Breed. 2017;77:185–98. <https://doi.org/10.5958/0975-6906.2017.00026.8>.
12. Gupta A, Kumar V, Kumar P, Pal R, Gurudyal, Singh C, Tyagi BS, Singh G. (2023) Compendium of wheat varieties notified in India during 2018–2023. Research Bulletin 51, ICAR- Indian Institute of Wheat & Barley Research, Karnal- 132001, India, pp 32.
13. Tanno K, Willcox G. How fast was wild wheat domesticated? Science. 2006;311:1886. <https://doi.org/10.1126/science.1124635>.
14. He XY, Wang JW, Ammar K, Peña RJ, Xia XC, He ZH. Allelic variants at the *Psy-A1* and *Psy-B1* loci in durum wheat and their associations with grain yellowness. Crop Sci. 2009;49:6. <https://doi.org/10.2135/cropsci2008.11.0651>.
15. Pont C, Leroy T, Seidel M, Alessandro T, Wandrille D, David A, Lang D, Bustos-Korts D, Goué N, Balfourier F, Molnár-Láng M, Lage J, Kilian B, Özkan H, Waite D, Dyer S, Letellier T, Alaux M, Consortium WHEALBI, Russell J, Keller B, van Eeuwijk F, Spannagl M, Mayer KFX, Waugh R, Stein N, Cattivelli L, Haberer G, Charmet G, Salse J. Tracing the ancestry of modern bread wheats. Nat Genet. 2019;51:905–11. <https://doi.org/10.1038/s41588-019-0393-z>.
16. Jordan KW, Wang S, Lun Y, Gardiner L-J, MacLachlan R, Hucl P, Wiebe K, Wong D, Forrest KL, Consortium IWGS, Sharpe AG, Sidebottom CHD, Hall N, Toomajian C, Close T, Dubcovsky J, Akhunova A, Talbert L, Bansal UK, Bariana HS, Hayden MJ, Pozniak C, Jeddalo JA, Hall A, Akhunov E. A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. Genome Biol. 2015;16:48. <https://doi.org/10.1186/s13059-015-0606-4>.
17. Levy AA, Feldman M. Evolution and origin of bread wheat. Plant Cell. 2022;34(7):2549–67. <https://doi.org/10.1093/plcell/koac130>.
18. Yan L, Helguera M, Kato K, Fukuyama S, Sherman J, Dubcovsky J. Allelic variation at the *VRN-1* promoter region in polyploid wheat. Theor Appl Genet. 2004;109(8):1677–86. <https://doi.org/10.1007/s00122-004-1796-4>.
19. Fu DL, Szucs P, Yan LL, Helguera M, Skinner JS, von Zitzewitz J, Hayes PM, Dubcovsky J. Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. Mol Genet Genomics. 2005;273:54–65. <https://doi.org/10.1007/s00438-004-1095-4>.
20. Milec Z, Tomková L, Sumíková T, Pánková K. A new multi plex PCR test for the determination of *Vrn-B1* alleles in bread wheat (*Triticum aestivum* L.). Mol Breed. 2012;30:317–23. <https://doi.org/10.1007/s11032-011-9621-7>.
21. Beales J, Turner A, Griffiths S, Snape JW, Laurie DA. A pseudo-response regulator is misexpressed in the photoperiod-insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). Theor Appl Genet. 2007;115(5):721–33. <https://doi.org/10.1007/s00122-007-0603-4>.
22. Ellis MH, Spielmeier W, Gale KR, Rebetzke GJ, Richards RA. Perfect markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. Theor Appl Genet. 2002;105:1038–42. <https://doi.org/10.1007/s00122-002-1048-4>.

23. Yang Y, Ma YZ, Xu ZS, Chen XM, He ZH, Yu Z, Wilkinson M, Jones HD, Shewry PR, Xia LQ. Isolation and characterization of Viviparous-1 genes in wheat cultivars with distinct ABA sensitivity and pre-harvest sprouting tolerance. *J Exp Bot.* 2007;58(11):2863–71. <https://doi.org/10.1093/jxb/erm073>.
24. Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T, Handa H, Ishida H, Mori M, Kawaura K, Ogiwara Y, Miura H. A wheat homolog of mother of FT and TFL1 acts in the regulation of germination. *Plant Cell.* 2011;23(9):3215–29. <https://doi.org/10.1105/tpc.111.088492>.
25. Lei L, Zhu X, Wang S, Zhu M, Carver BF, Yan L. *TaMFT-A1* is associated with seed germination sensitive to temperature in winter wheat. *PLoS ONE.* 2013;8(9):e73330. <https://doi.org/10.1371/journal.pone.0073330>.
26. Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G. Cloning and characterization of a critical regulator for preharvest sprouting in wheat. *Genetics.* 2013;195(1):263–73. <https://doi.org/10.1534/genetics.113.152330>.
27. Zhang YJ, Miao XL, Xia XC, He ZH. Cloning of seed dormancy genes (*TaSdr*) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. *Theor Appl Genet.* 2014;127:855–66. <https://doi.org/10.1007/s00122-014-2262-6>.
28. Liu G, Liu D, Zhang A, Liu H, Mia MS, Mullan D, Yan G. Identification of KASP markers and candidate genes for drought tolerance in wheat using 90K SNP array genotyping of near-isogenic lines targeting a 4BS quantitative trait locus. *Theor Appl Genet.* 2023;136:190. <https://doi.org/10.1007/s00122-023-04438-3>.
29. Jiang Q, Hou J, Hao C, Wang L, Ge H, Dong Y, Zhang X. The wheat (*T. aestivum*) sucrose synthase 2 gene (*TaSus2*) active in endosperm development is associated with yield traits. *Funct Integr Genomics.* 2011;11:49–61. <https://doi.org/10.1007/s10142-010-0188-x>.
30. Dong LL, Wang FM, Liu T, Dong ZY, Li AL, Jing RL, Mao L, Li YW, Liu X, Zhang KP, Wang DW. Natural variation of *TaGASR7-A1* affects grain length in common wheat under multiple cultivation conditions. *Mol Breed.* 2014;34:937–47. <https://doi.org/10.1007/s11032-014-0087-2>.
31. Hou J, Jiang Q, Hao C, Wang Y, Zhang H, Zhang X. Global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiol.* 2014;164(4):1918–29. <https://doi.org/10.1104/pp.113.232454>.
32. Qin L, Hao C, Hou J, Wang Y, Li T, Wang L, Ma Z, Zhang X. Homologous haplotypes, expression, genetic effects and geographic distribution of the wheat yield gene *TaGW2*. *BMC Plant Biol.* 2014;14:107. <https://doi.org/10.1186/1471-2229-14-107>.
33. Zhang YJ, Liu JD, Xia XC, He ZH. *TaGS-D1*, an ortholog of rice *OsGS3*, is associated with grain weight and grain length in common wheat. *Mol Breed.* 2014;34:1097–107. <https://doi.org/10.1007/s11032-014-0102-7>.
34. Zhang B, Liu X, Xu W, Chang JZ, Li A, Mao XG, Zhang XY, Jing RL. Novel function of a putative *MOC1* ortholog associated with spikelet number per spike in common wheat. *Sci Rep.* 2015;5:12211. <https://doi.org/10.1038/srep12211>.
35. Irshad A, Guo H, Ur Rehman S, Wang X, Gu J, Xiong H, Xie Y, Zhao L, Zhao S, Wang C, Liu L. Identification of single nucleotide polymorphism in *TaSBEIII* and development of KASP marker associated with grain weight in wheat. *Front Genet.* 2021;12. <https://doi.org/10.3389/fgene.2021.697294>.
36. Zhang W, Zhao J, He J, Kang L, Wang X, Zhang F, Hao C, Ma X, Chen D. Functional gene assessment of bread wheat: breeding implications in Ningxia Province. *BMC Plant Biol.* 2021;21:103. <https://doi.org/10.1186/s12870-021-02870-5>.
37. Liu SX, Chao S, Anderson J. New DNA markers for high molecular weight glutenin subunits in wheat. *Theor Appl Genet.* 2008;118:177–83. <https://doi.org/10.1007/s00122-008-0886-0>.
38. Ragupathy R, Naeem HA, Reimer E, Lukow OM, Sapirstein HD, Cloutier S. Evolutionary origin of the segmental duplication encompassing the wheat *GLU-B1* locus encoding the overexpressed Bx7 (Bx7^{OE}) high molecular weight glutenin subunit. *Theor Appl Genet.* 2008;116:283–96. <https://doi.org/10.1007/s00122-007-0666-2>.
39. Giroux MJ, Morris CF. A glycine to serine change in puroindoline b is associated with wheat grain hardness and low levels of starch-surface friabilin. *Theor Appl Genet.* 1997;95:857–64. <https://doi.org/10.1007/s001220050636>.
40. Chen F, Beecher BS, Morris CF. Physical mapping and a new variant of *Puroindoline b-2* genes in wheat. *Theor Appl Genet.* 2010;120:745–51. <https://doi.org/10.1007/s00122-009-1195-y>.
41. Nakamura T, Vrinten P, Saito M, Konda M. Rapid classification of partial waxy wheats using PCR-based markers. *Genome.* 2002;45(6):1150–6. <https://doi.org/10.1139/g02-090>.
42. Cha J-K, Park H, Kwon Y, Lee S-M, Oh K-W, Lee J-H. Genotyping the high protein content gene *NAM-B1* in wheat (*Triticum aestivum* L.) and the development of a KASP marker to identify a functional haplotype. *Agronomy.* 2023;13(8):1977. <https://doi.org/10.3390/agronomy13081977>.
43. Jiang Y, Jiang Q, Hao C, Hou J, Wang L, Zhang H, Zhang S, Chen X, Zhang X. A yield-associated gene *TaCWI*, in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis. *Theor Appl Genet.* 2015;128:131–43. <https://doi.org/10.1007/s00122-014-2417-5>.
44. Hanif M, Gao FM, Liu JD, Wen W, Zhang Y, Rasheed A, Xia XC, He ZH, Cao S. *TaTGW6-A1*, an ortholog of rice *TGW6*, is associated with grain weight and yield in bread wheat. *Mol Breed.* 2016;36:1. <https://doi.org/10.1007/s11032-015-0425-z>.
45. Wei B, Jing RL, Wang CS, Chen JB, Mao XG, Chang XP, Jia JZ. *Dreb1* genes in wheat (*Triticum aestivum* L.): development of functional markers and gene mapping based on SNPs. *Mol Breed.* 2009;23:13–22. <https://doi.org/10.1007/s11032-008-9209-z>.
46. Zhang J, Xu Y, Chen W, Dell B, Vergauwen R, Biddulph B, Khan N, Luo H, Appels R, Van den Ende W. A wheat *1-FEH w3* variant underlies enzyme activity for stem WSC remobilization to grain under drought. *New Phytol.* 2015;205(1):293–305. <https://doi.org/10.1111/nph.13030>.
47. Bernardo AN, Ma HX, Zhang D, Bai GH. Single nucleotide polymorphism in wheat chromosome region harboring *Fhb1* for Fusarium head blight resistance. *Mol Breed.* 2012;29:477–88. <https://doi.org/10.1007/s11032-011-9565-y>.
48. Singh L, Anderson JA, Chen J, Gill BS, Tiwari VK, Rawat N. Development and validation of a perfect KASP marker for Fusarium head blight resistance gene *Fhb1* in wheat. *Plant Pathol J.* 2019;35(3):200–7. <https://doi.org/10.5423/PPJ.OA.01.2019.0018>.

49. Neelam K, Brown-Guedira G, Huang L. (2013) Development and validation of a breeder-friendly KASPar marker for wheat leaf rust resistance locus *Lr21*. *Mol Breeding* 31:233–237. <https://doi.org/10.1007/s11032-012-9773-0>
50. Herrera-Foessel SA, Singh RP, Huerta-Espino J, Rosewarne GM, Periyannan SK, Viccar L, Calvo-Salazar V, Lan C, Lagudah ES. *Lr68*: a new gene conferring slow rusting resistance to leaf rust in wheat. *Theor Appl Genet.* 2012;124:1475–86. <https://doi.org/10.1007/s00122-012-1802-1>.
51. Terracciano I, Maccaferri M, Bassi F, Mantovani P, Sanguineti MC, Salvi S, Simkova H, Dolezel J, Massi A, Ammar K, Kolmer J, Tuberosa R. Development of COS-SNP and HRM markers for high-throughput and reliable haplotype-based detection of *Lr14a* in durum wheat (*Triticum durum* Desf). *Theor Appl Genet.* 2013;126:1077–101. <https://doi.org/10.1007/s00122-012-2038-9>.
52. Ramirez-Gonzalez RH, Segovia V, Bird N, Fenwick P, Holdgate S, Berry S, Jack P, Caccamo M, Uauy C. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnol J.* 2015;13(5):613–24. <https://doi.org/10.1111/pbi.12281>.
53. Zeller F. (1973) 1B/1R Wheat-rye chromosome substitutions and translocations. In: Sears ER, Sears LMS, editors *Proceedings of the 4th international wheat genetics symposium*, University of Missouri, Columbia, pp 209–222.
54. Varshney RK, Graner A, Sorrells ME. Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* 2005;10:621–30. <https://doi.org/10.1016/j.tplants.2005.10.004>.
55. Bagge M, Xia XC, Lu J, Berstedt T. Functional markers in wheat. *Curr Opin Plant Biol.* 2007;10:211–6. <https://doi.org/10.1016/j.pbi.2007.01.009>.
56. Rasheed A, Wen W, Gao F, Zhai S, Jin H, Liu J, Guo Q, Zhang Y, Dreisigacker S, Xia X, He Z. Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor Appl Genet.* 2016;129(10):1843–60. <https://doi.org/10.1007/s00122-016-2743-x>.
57. Hickey LT, Hafeez AN, Robinson H, Jackson SA, Leal-Bertioli SCM, Tester M, Gao C, Godwinn ID, Hayes BJ, Wulff BBH. Breeding crops to feed 10 billion. *Nat Biotechnol.* 2019;37:744–54. <https://doi.org/10.1038/s41587-019-0152-9>.
58. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 1980;8(19):4321–6. <https://doi.org/10.1093/nar/8.19.4321>.
59. Khalid M, Afzal F, Gul A, Amir R, Subhani A, Ahmed Z, Mahmood Z, Xia X, Rasheed A, Zhonghu H. Molecular characterization of 87 functional genes in wheat diversity panel and their association with phenotypes under well-watered and water-limited conditions. *Front Plant Sci.* 2019;10:717. <https://doi.org/10.3389/fpls.2019.00717>.
60. Zhao J, Wang Z, Liu H, Zhao J, Li T, Hou J, Zhang X, Hao C. Global status of 47 major wheat loci controlling yield, quality, adaptation and stress resistance selected over the last century. *BMC Plant Biol.* 2019;19:1–14. <https://doi.org/10.1186/s12870-018-1612-y>.
61. Perrier X, Jacquemoud-Collet JP. (2006) DARwin Software. <http://darwin.cirad.fr/darwin>
62. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 2010;11:94. <https://doi.org/10.1186/1471-2156-11-94>.
63. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. *Am J Hum Genet.* 2000;67:170–81. <https://doi.org/10.1086/302959>.
64. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* 2005;14:2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
65. Liu K, Muse SV. PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics.* 2005;21:2128–9. <http://dx.doi.org/10.1093/bioinformatics/bti282>.
66. Kumar A, Singh R, Singh G, Sharma RK, Saharan MS, Chhokar RS, Tyagi BS, Sendhil R, Chand R, Sharma I. Wheat cultivation in India (Pocket Guide). Directorate of Wheat Research, Karnal-132001. *Ext Bull.* 2014;52:1–34.
67. Singh G, Tyagi BS, Gupta A, Kumar V, Tiwari V, Chatrath R, Tiwari R, Singh SK, Saharan MS, Sharma RK, Gupta RK, Sharma I. Wheat: A guide on special features of varieties for different production conditions in India. Karnal: Directorate of Wheat Research; 2016.
68. Li C, Zhuang L, Li T, Hou J, Liu H, Jian C, Li H, Zhao J, Liu Y, Xi W, Hao P, Liu S, Si X, Wang X, Zhang X, Hao C. Conservatively transmitted alleles of key agronomic genes provide insights into the genetic basis of founder parents in bread wheat (*Triticum aestivum* L). *BMC Plant Biol.* 2023;23:100. <https://doi.org/10.1186/s12870-023-04098-x>.
69. Babalola KO, Monacelli N, Gozzi M, Ceccarelli S, Folloni S, Galaverna G. Genetic diversity and climate change adaptation in wheat: a systematic review of landraces, composite cross populations, and evolutionary populations. *Front Sustain Food Syst.* 2025;9:1504922. <https://doi.org/10.3389/fsufs.2025.1504922>.
70. Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Are AK, Ortiz R. Landrace germplasm for improving yield and abiotic stress adaptation. *Trends Plant Sci.* 2016;21(1):31–42. <https://doi.org/10.1016/j.tplants.2015.10.012>.
71. Jobson EM, Johnston RE, Oiestad AJ, Martin JM, Giroux MJ. The impact of the wheat *Rht-B1b* semi-dwarfing allele on photosynthesis and seed development under field conditions. *Front Plant Sci.* 2019;10:51. <https://doi.org/10.3389/fpls.2019.00051>.
72. Quagliata G, Maghrebi M, Marín-Sanz M, Palombieri S, Sestili F, Lafiandra D, Barro F, Vigani G, Astolfi S. Rye-durum wheat 1BL.1RS translocation: implications for drought tolerance and nutritional status. *Physiol Plant.* 2024;176(5):e14579. <https://doi.org/10.1111/ppl.14579>.
73. Li L, Sun Z, Zhang Y, Ke H, Yang J, Li Z, Wu L, Zhang G, Wang X, Ma Z. Development and utilization of functional Kompetitive Allele-Specific PCR markers for key genes underpinning fiber length and strength in *Gossypium hirsutum* L. *Front Plant Sci.* 2022;13:853827. <https://doi.org/10.3389/fpls.2022.853827>.
74. Khan MK, Pandey A, Thomas G, Akkaya MS, Kayis SA, Ozsensoy Y, Hamurcu M, Gezgin S, Topal A, Hakki EE. Genetic diversity and population structure of wheat in India and Turkey. *AoB Plants.* 2015;7:plv083. <https://doi.org/10.1093/aobpla/plv083>.

75. Cheng S, Feng C, Wingen LU, Cheng H, Riche AB, Jiang M, Leverington-Waite M, Huang Z, Collier S, Orford S, Wang X, Awal R, Barker G, O'Hara T, Lister C, Siluveru A, Quiroz-Chávez J, Ramírez-González RH, Bryant R, Berry S, Bansal U, Bariana HS, Bennett MJ, Bicego B, Bilham L, Brown JKM, Burridge A, Burt C, Buurman M, Castle M, Chartrain L, Chen B, Denbel W, Elkot AF, Fenwick P, Feuerhelm D, Foulkes J, Gaju O, Gauley A, Gaurav K, Hafeez AN, Han R, Horler R, Hou J, Iqbal MS, Kerton M, Kondic-Spica A, Kowalski A, Lage J, Li X, Liu H, Liu S, Lovegrove A, Ma L, Mumford C, Parmar S, Philp C, Playford D, Przewieslik-Allen AM, Sarfraz Z, Schafer D, Shewry PR, Shi Y, Slafer GA, Song B, Song B, Steele D, Steuernagel B, Tailby P, Tyrrell S, Waheed A, Wamalwa MN, Wang X, Wei Y, Winfield M, Wu S, Wu Y, Wulff BBH, Xian W, Xu Y, Xu Y, Yuan Q, Zhang X, Edwards KJ, Dixon L, Nicholson P, Chayut N, Hawkesford MJ, Uauy C, Sanders D, Huang S, Griffiths S. Harnessing landrace diversity empowers wheat breeding. *Nature*. 2024;632:823–31. <https://doi.org/10.1038/s41586-024-07682-9>.
76. Singh G, Raigar OP, Kaur S, Bishnoi R, Mondal K, Abreha KB, Nayak AK, Athar T, Sharma V, Aziz D, Gudi S, Saini P, Kumar A, Bhardwaj R, Riar A. Advances in genomics-assisted breeding strategies for enhancing nutrient uptake and use efficiency in cereals: a pathway toward sustainable agriculture. *Plant Stress*. 2025;18:101002. <https://doi.org/10.1016/j.stress.2025.101002>.
77. Sun C, Hu H, Cheng Y, Yang X, Qiao Q, Wang C, Zhang L, Chen D, Zhao S, Dong Z, Chen F. Genomics-assisted breeding: the next-generation wheat breeding era. *Plant Breed*. 2023;142(3):259–68. <https://doi.org/10.1111/pbr.13094>.
78. Scott MF, Ladejobi O, Amer S, Bentley AR, Biernaskie J, Boden SA, Clark M, Dell'Acqua M, Dixon LE, Filippi CV, Fradgley N, Gardner KA, Mackay IJ, O'Sullivan D, Percival-Alwyn L, Roorkiwal M, Singh RK, Thudi M, Varshney RK, Venturini L, Whan A, Cockram J, Mott R. Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity (Edinb)*. 2020;125(6):396–416. <https://doi.org/10.1038/s41437-020-0336-6>.
79. Mao H, Jiang C, Tang C, Nie X, Du L, Liu Y, Cheng P, Wu Y, Liu H, Kang Z, Wang X. Wheat adaptation to environmental stresses under climate change: molecular basis and genetic improvement. *Mol Plant*. 2023;16:1564–89. <https://doi.org/10.1016/j.molp.2023.09.001>.
80. Fourquet L, Barber T, Campos-Mantello C, Howell P, Orman-Ligeza B, Percival-Alwyn L, Rose GA, Sheehan H, Wright TIC, Longin F, Würschum T, Novoselovic D, Greenland AJ, Mackay IJ, Cockram J, Bentley AR. An eight-founder wheat MAGIC population allows fine-mapping of flowering time loci and provides novel insights into the genetic control of flowering time. *Theor Appl Genet*. 2024;137(12):277. <https://doi.org/10.1007/s00122-024-04787-7>.
81. Aoun M, Carter AH, Morris CF, Kiszonas AM. Genetic architecture of end-use quality traits in soft white winter wheat. *BMC Genomics*. 2022;23:440. <https://doi.org/10.1186/s12864-022-08676-5>.
82. Cao S, Xu D, Hanif M, Xia X, He Z. Genetic architecture underpinning yield component traits in wheat. *Theor Appl Genet*. 2020;133:1831–43. <https://doi.org/10.1007/s00122-020-03562-8>.
83. Bevan MW, Uauy C, Wulff BB, Zhou J, Krasileva K, Clark MD. Genomic innovation for crop improvement. *Nature*. 2017;543(7645):346–54. <https://doi.org/10.1038/nature22011>.
84. Ennos RA. Maintenance of genetic variation in plant populations. In: Hecht MK, Wallace B, Prance GT, editors. *Evolutionary Biology*. Boston, MA: Springer; 1983. https://doi.org/10.1007/978-1-4615-6971-8_4.
85. Goel S, Rathore M, Grewal S, Jain N, Singh BK, Ahlawat AK, Singh AM, Singh PK, Singh NK. Effect of allelic variation in triticin on bread- and chapati-making qualities of wheat (*Triticum aestivum*). *Agric Res*. 2015;4:139–51. <https://doi.org/10.1007/s40003-015-0150-1>.
86. Plavšin I, Gunjača J, Šatović Z, Šarčević H, Ivić M, Dvojković K, Novoselović D. An overview of key factors affecting genomic selection for wheat quality traits. *Plants*. 2021;10(4):745. <https://doi.org/10.3390/plants10040745>.

Figures

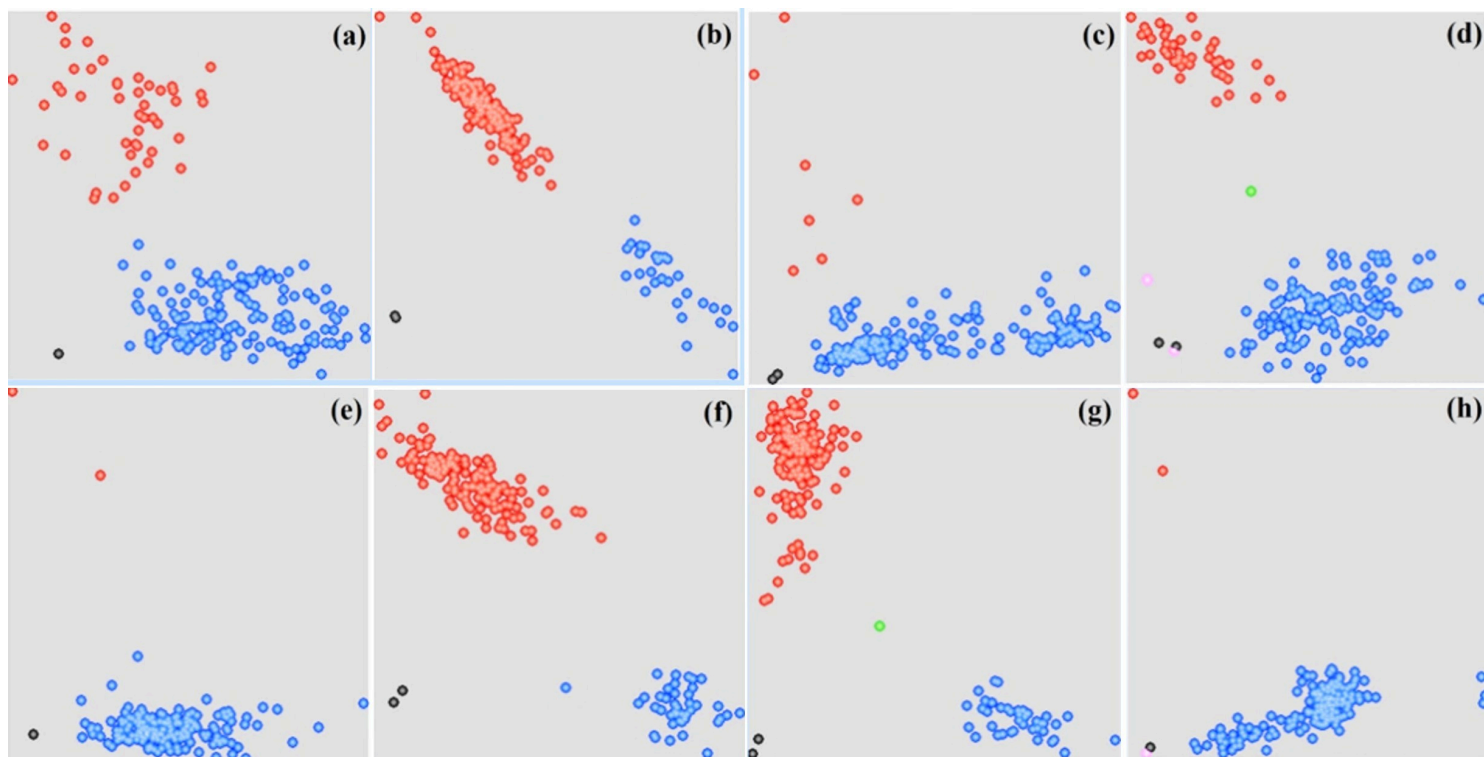


Figure 1

Scatter plots for selected KASP assays show clustering of wheat genotypes along the X (FAM) and Y (HEX) axes. Genotypes colored blue have the FAM-type allele, genotypes colored red have the HEX-type allele, and black dots represent the non-template control (NTC). (a) KASP assay for *DreB-B1* locus shows *DreB-B1a* allele on the FAM axis and *DreB-B1b* allele on the HEX axis, (b) KASP assay for *Glu-A1* locus shows *Ax1* or *Ax2** allele on the FAM axis and *Null* allele on the HEX axis, (c) KASP assay for *GS-D1* locus shows the *GS-D1a* allele on the FAM axis and the *GS-D1b* allele on the HEX axis, (d) KASP assay for *Lr34* locus shows the *Lr34+* allele on the FAM axis and the *Lr34-* allele on the HEX axis, (e) KASP assay for *Fhb1* locus shows the *Fhb1_R* allele on the FAM axis and the *Fhb1_S3* allele on the HEX axis, (f) KASP assay for *Pina-D1* locus shows the *Pina-D1a* allele on the FAM axis and the *Pina-D1b* allele on the HEX axis. (g) KASP assay for *Ppd-D1* locus shows the *Ppd-D1a* allele on the FAM axis and the *Ppd-D1b* allele on the HEX axis. (h) KASP assay for *Yr15* locus shows the *R5_Yr15-* allele on the FAM axis and the *R5_Yr15+* allele on the HEX axis.

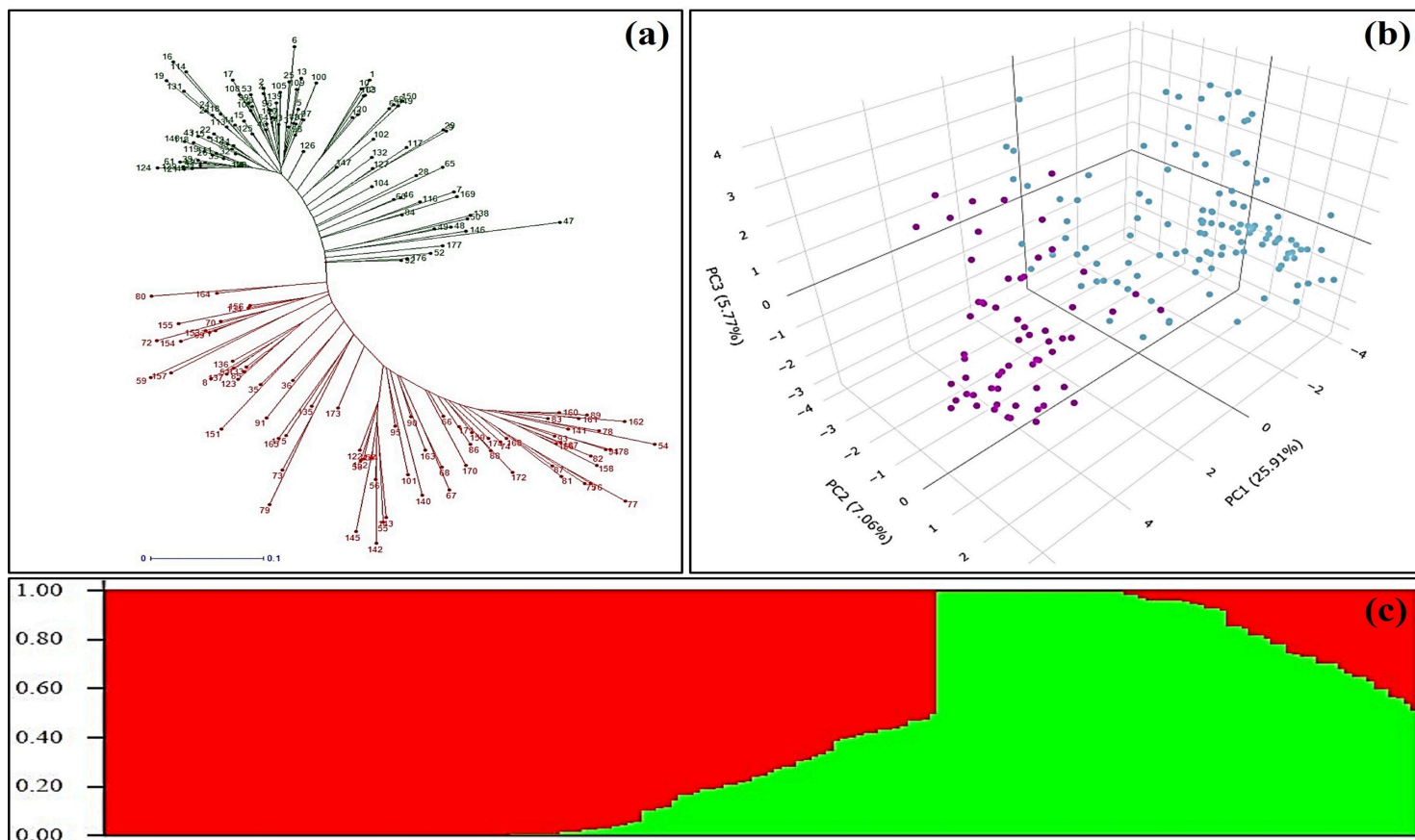


Figure 2

Population structure analysis of 178 wheat accessions using 40 functional KASP marker loci. (a) A neighbour-joining (NJ) phylogenetic tree illustrating the genetic relationships among the accessions, constructed based on pairwise genetic distances derived from the KASP marker data. Branch lengths represent genetic dissimilarity, and accessions are grouped according to genetic similarity. (b) Principal component analysis (PCA) of the accessions, showing the distribution of genetic variation along the first three principal components (PC1, PC2, and PC3). Each point represents an individual accession, and spatial clustering reflects underlying genetic structure. (c) Model-based population structure analysis using the STRUCTURE software, with the optimal number of subpopulations (K) inferred based on the admixture model.

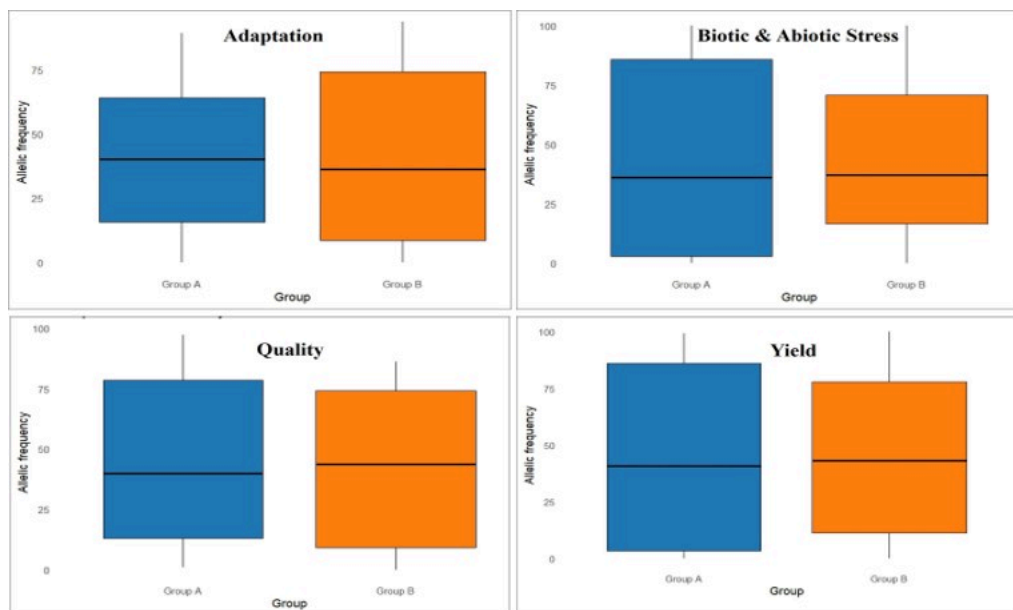


Figure 3

Allelic frequency distribution of functional gene categories across two STRUcTURE-defined wheat subpopulations (Group A and Group B). Box plots illustrate comparable allelic diversity for adaptation, biotic–abiotic stress, quality, and yield traits, with only minor median shifts indicating minor subgroup

differentiation under broadly shared selection pressures.

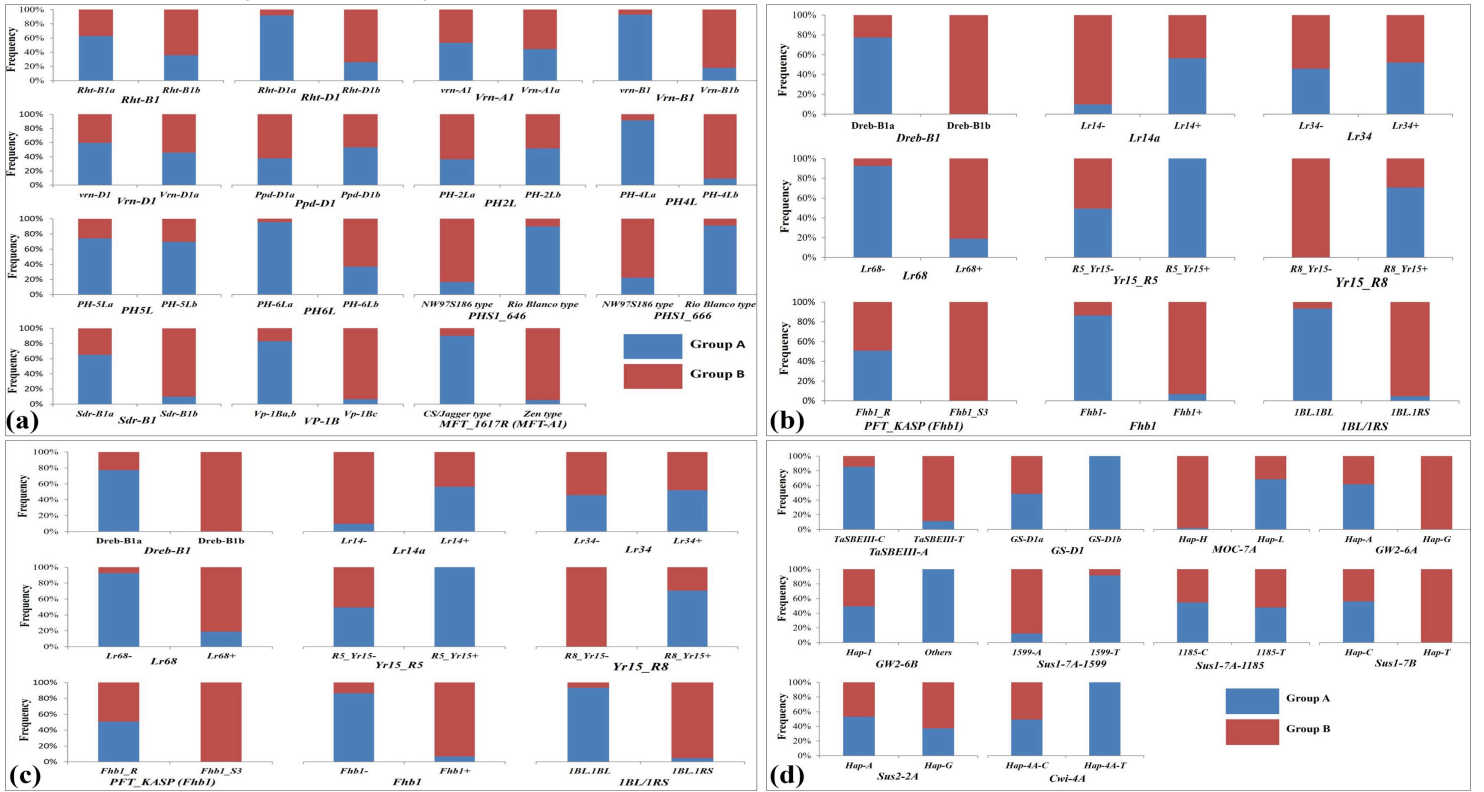


Figure 4

Allelic frequency distribution of genes between two STRUCTURE-defined subpopulations (Group A and Group B). The stacked bar plots illustrate the proportion of alleles across different gene categories. (a) adaptation-related genes, (b) biotic and abiotic stress-related genes, (c) grain quality-related genes, and (d) yield-related genes. The red and blue colours represent Group B and Group A, respectively, showing differential enrichment patterns across loci and trait categories.

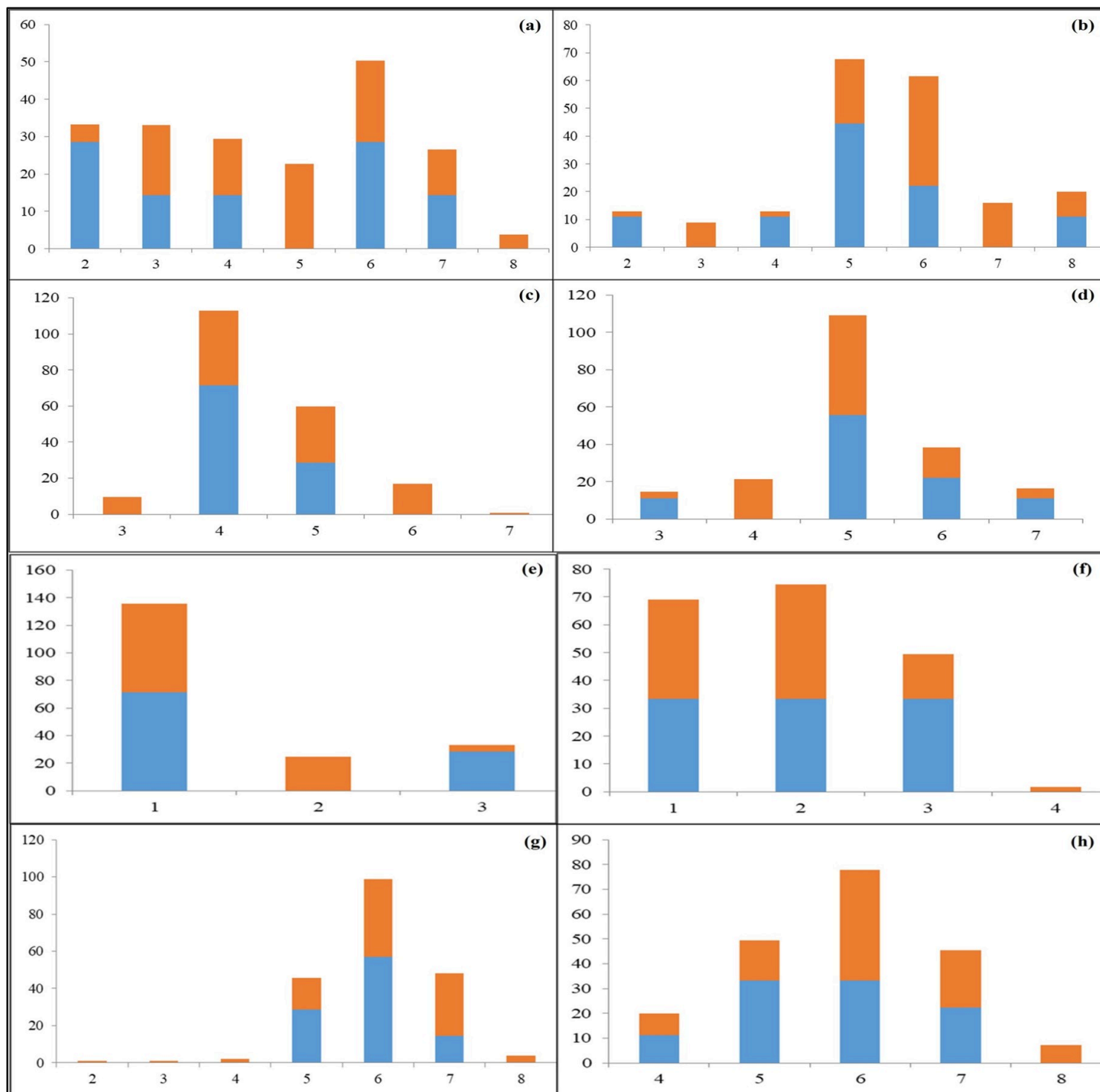


Figure 5

Distribution of favourable alleles associated with adaptation, biotic and abiotic stress tolerance, grain quality, and yield-related traits among the seven most frequently used parental lines and the remaining 171 bread wheat genotypes constituting two STRUCTURE-defined subpopulations. **(a, b)** Frequency distribution of adaptation-related favourable alleles in Group A and Group B subpopulations, respectively; **(c, d)** biotic and abiotic stress-related favourable alleles in Group A and Group B; **(e, f)** quality-related favourable alleles in Group A and Group B; and **(g, h)** yield-related favourable alleles in Group A and Group B. The X-axis represents the number of favourable alleles, and the Y-axis represents the frequency of genotypes within each class. Blue bars indicate the seven most frequently used founder parents, while brown bars represent the derived or remaining panel genotypes.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table3.docx](#)

- [Table1.docx](#)
- [Table2.docx](#)
- [SupplementaryTableS1.xlsx](#)
- [SupplementaryTableS2.docx](#)