

# Diversity in Indian wheat (*T. aestivum* L.) germplasm for various agro-morphological traits

Sabhyata Sabhyata<sup>1,2</sup>, Arun Gupta<sup>1</sup> , Diwakar Aggarwal<sup>2</sup>, Ratan Tiwari<sup>1</sup>,  
Gyanendra Singh<sup>1</sup> and Gyanendra Pratap Singh<sup>3</sup>

## Research Article

**Cite this article:** Sabhyata S, Gupta A, Aggarwal D, Tiwari R, Singh G, Singh GP (2023). Diversity in Indian wheat (*T. aestivum* L.) germplasm for various agro-morphological traits. *Plant Genetic Resources: Characterization and Utilization* **21**, 81–89. <https://doi.org/10.1017/S1479262123000461>

Received: 14 December 2022  
Revised: 24 June 2023  
Accepted: 26 June 2023  
First published online: 21 July 2023

### Keywords:

Cluster analysis; diversity; landraces; PCA; *T. aestivum*

### Corresponding author:

Arun Gupta;  
Email: [arung66@yahoo.com](mailto:arung66@yahoo.com)

<sup>1</sup>ICAR-Indian Institute of Wheat and Barley Research, Karnal-132001, Haryana, India; <sup>2</sup>Department of Biotechnology, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala-133207, Haryana, India and <sup>3</sup>ICAR-National Bureau of Plant Genetic Resources, New Delhi, India

### Abstract

A better understanding of diversity in landraces is essential for planning crosses for the development of trait specific varieties with better adaptability and stability. In the present study, 120 wheat genotypes comprised of landraces, genetic stocks, released varieties and improved genotypes were evaluated in randomized block design for 14 agro-morphological traits. The clustering method and principal component analysis (PCA) programme of Statistical Package for Agricultural Research (SPAR1) was used for grouping the genotypes. These 120 genotypes were grouped into nine clusters based on agro-morphological traits. These nine clusters differed significantly on the basis of mean values for 14 agro-morphological traits. PCA showed that the two principal components (PC1 to PC2) exhibited about 49% of the total variability. Scatter plot was constructed by plotting scores of PC1 and PC2. Based on mean values obtained over two years, the diverse superior genotypes were identified for utilization in hybridization programme. From the present study, we conclude that cluster analysis grouped the landraces with greater agro-morphological similarity into one group rather than geographical isolation indicating that geographical origin may not be the only factor causing diversity. Further, released varieties exhibited superiority for grain yield, while many landraces had higher values for number of tillers in a meter, biomass and thousand-grain weight (TGW). Thus, for the improvement in TGW in released varieties, the hybridization between superior landraces for TGW from cluster 'E' and released wheat varieties from cluster 'C' could give desirable segregates.

## Introduction

Wheat (*Triticum* spp.) is a cosmopolitan crop due to its cultivation in all the major agro-ecological zones of the world. It is the second most important cereal crop after rice and accounts for about one-third of global grains production (Tandon, 2000). Ninety per cent of the world's wheat acreage is occupied by *T. aestivum* ssp *aestivum* L, while remaining comes from other two species *T. turgidum* ssp *durum* (Desf.) Husn. and *T. turgidum* ssp *dicoccum* (Schrank ex Schubl) Thell. (Priya *et al.*, 2015). During the course of evolution, farmers selected array of characters such as plant height (PH), higher number of spike and grains, spike weight and 1000-grain weight (Peng *et al.*, 2011). Local landraces evolved in diverse ecosystem are important source of genetic variation, which greatly contributes to the improvement of wheat varieties (Simmonds, 1979). Many landraces from India have been identified for specific traits, such as *Muzaffarnagar white* for high yielding, *Mundia* for drought tolerance, *Halna* for heat tolerance, *Naphal* suitable for biscuit making. The wheat landrace *Daruma* from Japan is a source of reduced PH in majority of modern cultivars (Reitz and Salmon, 1968). The landraces are adapted to specific niche and having specific attributes preferred by the farmers. Genetic variation in the experimental material provides an information base of traits to the breeders for utilization in hybridization programme. Studying the diversity among the genotypes, landraces and varieties is useful for improving and increasing genetic variation, and to develop strategies for the conservation and management of genetic resources and its utilization in breeding programmes (Priya *et al.*, 2015). Therefore, information about diversity within the germplasm collection has a significant impact on crop improvement and is essential for the development of eco-friendly, disease resistant, high yielding wheat varieties suitable for diverse climatic conditions (Wani *et al.*, 2018). Even for developing hybrid wheat, requires complete understanding of diversity in floral structure (Maniee *et al.*, 2009). Landraces are important genetic resources that can be used in development of high yielding varieties especially under changing climatic condition (Jaradat, 2013). Further, they have a closer association with new cultivars and main constitution of primary gene pool thus can be used



for widening the genetic base of the wheat crop especially under changing climatic condition. To study the genetic relationship among the genotypes multivariate statistical algorithms is an important approach (Mohammadi and Prasanna, 2003). Several researchers have proposed various multivariate statistical tools to study the diversity in wheat including the cluster analysis and principal component analysis (PCA) (Jamali and Ali, 2008; Degewione and Alamerew, 2013). Cluster analysis method can also be used for the assessment of valuable economic and adaptive traits (Ali *et al.*, 2019; Adilova *et al.*, 2020) and selection of high-performance genotypes. It was also used to determine the morphological similarities between genotypes (Awan *et al.*, 2014; Yadav *et al.*, 2015).

Further, in cluster analysis based on morphological similarities, the groups are formed and on the basis of genetic distance among groups, divergent genotypes are identified for hybridization programme. The hybridization among genetically different parents is likely to give higher heterotic expression in next-progeny (Shekhawat *et al.*, 2001). The degree of heterosis obtained in the offspring is proportional to the genetic distance between parents (Wani *et al.*, 2018). PCA can convert some potentially correlated variables into a small number of variables called principal components (Mujaju and Chakuya, 2008). In addition, the analysis describes the maximum variance in the data set by including the total variance of the variable and function as a linear variable (Mecha *et al.*, 2017).

The utilization of plant genetic resources has been emphasized by various authors (Dempewolf *et al.*, 2017; Hasan *et al.*, 2020) for the development of superior varieties and sustainable food production. Landraces are considered important from adaptability and genetic study point of view (Villa *et al.*, 2005). During the past two decades, characterization and evaluation of landraces are being considered as an important activity in any genetic resources project. The characterization of population structure (landraces/released cultivar) in wheat germplasm is essential to understand diversity. However limited studies are conducted in evaluation of landraces of wheat in India. The purpose of this study was to assess the diversity among released wheat varieties and wheat landraces collected from various parts of India by using cluster analysis and PCA and to identify the superior genotypes for utilization in hybridization programme.

## Materials and methodology

### Plant material

For the present study, 120 diverse wheat genotypes were obtained from germplasm resource unit of ICAR-Indian Institute of Wheat and Barley (ICAR-IIWBR), Karnal. Of the 120 wheat genotypes, 95 were landraces, 7 were trait specific registered genetic stocks, 14 were released and notified varieties, 3 were improved genotypes (Advance lines) and 1 was exotic material. Of the 95 landraces, 42 were collected from Gujarat, 24 from Madhya Pradesh, 12 from Uttarakhand, 7 from Rajasthan, 3 from Assam, 2 from Maharashtra and one each from Himachal Pradesh, Andhra Pradesh, Arunachal Pradesh, Delhi and Uttar Pradesh.

### Experimental site and design

The material was evaluated consecutively for two winter seasons during 2019–2020 and 2020–2021 at research and experimental farm of ICAR-IIWBR, Karnal (latitude 29°42' N and longitude 76°59'E) situated at an elevation of 240 m above sea level. The

experiment was carried out in randomized block design with two replications. Each genotype was grown in 1.25 m long three rows plot with row to row spacing of 25 cm between rows and 5–6 cm between plants. A basal dose of N:P:K @50:60:40 kg/ha was applied at the time of sowing. Another dose of Nitrogen@50 kg/ha was top dressed at tillering stage. Plots were irrigated at 20–25 days interval. Weeds were controlled manually. Propiconazole (25EC@0.1%) was sprayed at flag leaf opening stage (Zadoks growth stage 47) and kernel water ripe stage (Zadoks growth stage 71) to control yellow and brown rust. Crop was harvested manually by cutting above ground.

### Data recording

Data were recorded on 14 agro-morphological traits. Days to heading (DH), days to anthesis (DA), days to maturity (DM) and grain filling period (GFP) were recorded on plot basis. DH was recorded as the number of days taken from sowing to when ear emergence was observed in 50% of plants, DA was recorded as the number of days taken from sowing to when extrusion of anthers from the spikelets was observed in 50% plants, DM was recorded as number of days taken from sowing to when physiological maturity was observed in 75% plants showed and GFP was counted as number of days taken from anthesis to maturity. Five plants were randomly selected from the centre row and observations on PH, spike length (SL), spikelet number per spike (SIPS) and spike dry weight (SDW) were recorded from it. PH of the main tiller was measured in cm from ground level to the top of the spike excluding awn at maturity, SL was recorded in cm from the base of the spike to the tip excluding awns, SDW was weighed in gram after harvesting spike from main tiller, SIPS were counted as number of spikelets on a single spike of the main tiller, seed number per spike (SNS) were counted as number of seeds on a single spike and weighed for seed weight per spike (SWS). Number of tillers in a meter (TN) were counted and recorded; weight of randomly selected 1000 grains from each plot was recorded as thousand-grain weight (TGW). All the plants from each plot were harvested above the ground level and weight of total produce was recorded as Biomass (BM) and grains were threshed separately and weighed from each plot and recorded as grain yield (GY).

### Statistical analysis

Pooled analysis of variance was performed on replicated two years data recorded for each trait. The non-significant Genotype x Year effect was observed for all the traits except SNS, SWS, BM and SDW. As a result of this analysis, mean values from the two years data was calculated for each trait and used in further analysis. The clustering method and PCA programme of Statistical Package for Agricultural Research (SPAR1) was used for grouping the genotypes. Different cluster solutions were compared using a sequential F-ratio test as described by Beale (1969) and elaborated by Sparks (1973). Mean, standard deviation and coefficient of variation (Zhong and Qualset, 1995) were calculated for each group as well as for released varieties and landraces. R software was used for plotting graphs (R Core Team, 2020).

## Results

The mean values, range, standard deviation and coefficient of variation (CV%) for each trait as well as on overall mean basis

vis a vis for released varieties as a group vs landraces as a group were calculated and presented in Table 1. The highest CV values was observed for GY followed by SWS on over all mean basis as well as in released varieties and landraces groups. The lowest CV value was observed for DM indicating the least variation among all accessions for this trait. The released variety group exhibited significant superiority over landrace's group for GY and SWS, whereas landrace's group exhibited numerically higher value for PH, TN and BM as compared to released varieties.

### Grouping of genotypes

The cluster analysis program of SPAR1 grouped 120 genotypes into nine clusters namely cluster A, B, C, D, E, F, G, H and I. Maximum number of genotypes were found in cluster 'F' (30) followed by cluster 'I' (27 genotypes) while Cluster 'B' had only two genotypes. Cluster 'A' contained 15 genotypes mainly collection from Gujarat and Madhya Pradesh and registered genetic stocks. Cluster 'B' contained two genotypes namely Kapargaon local (*T. aestivum* ssp *sphaerococcum* (Perc.) Mac key) and Mangura local. Cluster 'C' had nine wheat genotypes mainly comprising of released varieties, registered genetic stocks. The collection from Uttarakhand was grouped in cluster 'D', while Cluster 'E' and 'G' contained eight and 15 genotypes, respectively and mainly collection from MP and Gujarat. Cluster 'F' comprised of landraces collected from Gujarat, Madhya Pradesh, Rajasthan, Uttarakhand, Assam and Maharashtra. Cluster 'H' comprised of landraces collected from Gujarat, Rajasthan, Uttarakhand, Assam and Madhya Pradesh. Cluster 'I' covered accessions collected from Gujarat, M.P., Rajasthan, Assam, Andhra Pradesh and Arunachal Pradesh.

### Intra cluster values and inter-relation of cluster

The inter-cluster distance varied from 2.218 to 8.65. The minimum inter-cluster distance was observed between cluster 'F' and cluster 'I' ( $d=2.218$ ) indicating that the genotypes of these two clusters were closely related. The maximum inter-cluster distance was observed among cluster 'C' and 'B' (8.65) followed by cluster 'D' and 'E' (8.35), cluster 'B' and 'H' (7.82), cluster 'B' and 'G' (7.75), cluster 'B' and 'E' (7.741), cluster 'A' and 'D' (7.387) (Table 2). Intra-cluster distance range varied from 2.39 to 3.02. The maximum intra-cluster distance was observed in cluster 'D' (3.024) followed by cluster 'E' (2.668) and cluster 'H' (2.667).

### Cluster mean analysis

Cluster mean and standard deviation of each trait were tabulated in table 3. Cluster 'B' comprised of only two accessions with 1.66% of the population and covering population with short PH ( $96.5 \pm 3.9$ ), lower value for SL ( $6.52 \pm 1.5$ ), lower value for SWS ( $1.02 \pm 0.02$ ) and TGW ( $24.99 \pm 5.5$ ), GY ( $135.6 \pm 24.6$ ), BM ( $0.63 \pm 0.1$ ), TN ( $64 \pm 2.1$ ), SDW ( $1.44 \pm 0.09$ ). Cluster 'C' had higher value for SPS ( $21 \pm 0.8$ ), SNPS ( $51 \pm 5.1$ ), SWS ( $2.2 \pm 0.29$ ), GY ( $386.7 \pm 55.3$ ) and SDW ( $2.94 \pm 0.3$ ). The cluster 'D' depicts 4.16% of the population and comprised five genotypes which showed higher value for DH ( $114 \pm 6.6$ ), DA ( $121 \pm 5$ ) and DM ( $153 \pm 3.4$ ) and lower value for GFP ( $39 \pm 5.7$ ). Cluster 'E' contributed 6.6% of the population with eight genotypes, having higher value for GFP ( $53 \pm 4$ ) and TGW ( $45.6 \pm 5.7$ ), while lower value for DH ( $87 \pm 3.4$ ), DA ( $98 \pm 2.6$ ) and lower value

for SPS ( $17 \pm 0.7$ ). Cluster 'F' has huge contribution i.e. 25% of the total population fall in this cluster and had lower value for SNS ( $35 \pm 4.1$ ). The cluster 'G' contained 15 genotypes and higher value for DM ( $139 \pm 3.6$ ) and SL ( $10.9 \pm 1.15$ ). Cluster 'H' contained 9 genotypes and higher value for BM ( $1.4 \pm 0.2$ ) and TN ( $94 \pm 6.8$ ). Cluster 'I' consist of 27 genotypes and had moderate mean values for all the traits.

The mean coefficient of variation in the nine clusters exposed that maximum per cent of variation was observed in cluster 'D' (12.32%) tailed by cluster 'B' (11.17%), 'H' (9.15%), 'E' (8.87%), 'F' (8.68%), 'I' (8.62%), 'G' (8.51%), 'A' (8.33%), 'C' (7.55%) as compared to the complete set (9.24%).

### Principal component analysis

The correlation between PCA score and 14 agro-morphological traits was calculated. The first two PC axes accounted for 49.18% of the total variance. PCA 1 explained 28.29% of the total variance and positively correlated to traits DH, DA, DM, PH and SDW, while negatively correlated with SL, SNS, SDW, GFP, TGW, BM, SPS, GY and TN. PCA 2 contributed to 20.89% of the total variation and depicted mainly positive correlation with traits DH, DA, DM, SL, SPS, SNS, SWS and SDW. PCA 2 showed a negative association with GFP and TGW only (Fig. 1(B)).

A Scatter plot between PC1 and PC2 depicted a pattern of genotypes in the factor plane and plotted distinct nine clusters in the present study (Fig. 1(A)). Relationship between all the traits using traits correlation plot is depicted in Fig. 1(B). The traits were superimposed on a principal component plot as vectors (Fig. 1(B)); the relative length of the vector represents the relative rate of variation in each represented variable.

The present study helped in identifying the superior genotypes for different traits (Table 4). For example, landraces LGM133, IC321892, IC321851, IC321910 and Girgehun taken less than 87days for DH. The released variety DBW187 was found superior for SPS, SNS, SWS, SDW, TGW, GY. A landrace IC321931 had various desirable attributes such as PH, SNS, SDW and SWS. Majority of superior genotypes for GFP, BM, TN and TGW are landraces. A heatmap of agro-morphological traits of the superior genotypes was constructed and revealed the level of diversity in superior genotypes. The colour richness related to diversity in different clusters (Fig. 2).

The mean values of 120 for 14 traits along with passport data are given in supplementary table S1.

### Discussion

The highly significant differences was observed for genotypes for all the studied traits in ANOVA indicate that there is wide variability exist in studied traits broadly and further suggest the ample scope of genetic manipulation. Further, the landraces are an important genetic material as they are the product of natural selection by domestication carried out by farmers over the century (Al Lawati *et al.*, 2021). They can easily adapt to adverse biotic and abiotic pressure. Large number of landraces of wheat was collected from 13 states (Andhra Pradesh, Arunachal Pradesh, Assam, Gujarat, Himachal Pradesh, Jammu and Kashmir, Karnataka, M.P., Maharashtra, Rajasthan, Sikkim, Tamil Nadu and Uttarakhand) in India (Panwar *et al.*, 2014). However, landraces have been replaced by improved varieties and no longer being cultivated in states like Punjab, Haryana and western U.P. This can be evident from the passport data of the studied material

**Table 1.** Descriptive statistics for 14 agro-morphological traits on an overall basis vis a vis released varieties versus landraces

Traits	Overall					Released varieties					Landraces				
	Range	Mean	SD	CV%		Range	Mean	SD	CV%		Range	Mean	SD	CV%	C.D. at 5%
DH (Days)	83.25-121.25	97.99	6.38	6.51		86.5-105.25	98.15	4.95	5.04		83.25-121.25	97.96	6.61	6.75	3.95
DA(Days)	95.5-126.75	106.05	5.73	5.40		98.5-115	105.69	4.90	4.64		95.5-126.75	106.10	5.88	5.54	4.48
DM (Days)	128-157.5	143.25	3.98	2.78		136.5-147.5	142.38	3.85	2.70		128-157.5	143.40	4.01	2.80	4.43
PH (cm)	84.5-132.25	108.88	8.98	8.25		84.5-120.88	101.29	7.95	7.85		90.75-132.25	110.13	8.55	7.76	11.41
SL (cm)	5.46-12.91	9.74	1.19	12.22		8.33-11.57	10.03	0.82	8.18		5.46-12.91	9.69	1.24	12.80	1.65
SIPS (no.)	16.5-21.75	18.7	1.26	6.74		16.75-21.75	19.44	1.52	7.82		16.5-21	18.58	1.18	6.35	2.47
SNS(no.)	25.7-60.62	40.92	6.65	16.25		34.15-57.3	45.06	7.09	15.73		25.7-60.62	40.24	6.35	15.78	10.23
SWS (g)	0.72-2.81	1.69	0.36	21.30		1.25-2.81	1.92	0.47	24.48		0.72-2.47	1.65	0.32	19.39	0.49
GFP(Days)	32.75-57.25	45.27	4.73	10.45		40.25-50	44.24	2.60	5.88		32.75-57.25	45.44	4.98	10.96	9.15
TGW (g)	20.44-55.06	38.6	5.11	13.24		34.3-45.17	39.09	2.92	7.47		20.44-55.06	38.52	5.39	13.99	1.08
GY (g)	114.28-531.75	279.27	81.5	29.18		150.89-479.7	349.62	99.60	28.49		114.28-531.75	267.66	72.33	27.02	26.54
BM (kg/plot)	0.56-2.08	1.2	0.21	17.50		0.97-1.58	1.19	0.17	14.29		0.56-2.08	1.20	0.22	18.33	0.26
TN (no.)	57.5-106.5	78.12	10.35	13.25		66.25-93	77.99	8.43	10.81		57.5-106.5	78.15	10.66	13.64	26.1
SDW (g)	1.15-3.63	2.24	0.45	20.09		1.56-3.63	2.57	0.63	24.51		1.15-3.17	2.19	0.39	17.81	0.38

**Table 2.** Intra and inter-cluster distance between nine clusters in wheat genotypes

Cluster	A	B	C	D	E	F	G	H	I
A	2.514								
B	7.165	2.465							
C	3.747	8.651	2.354						
D	7.387	6.836	7.03	3.024					
E	2.877	7.741	5.906	8.35	2.668				
F	3.505	5.308	5.66	5.424	3.756	2.423			
G	2.31	7.759	3.209	7.097	3.478	3.924	2.617		
H	4.901	7.821	4.185	5.165	5.874	4.058	3.757	2.667	
I	3.143	6.259	4.148	4.695	4.61	2.218	3.151	2.961	2.393

Note: intra cluster distance (diagonal and bold) and inter cluster distance (off-diagonal).

that among the indigenous collection, none of the landraces have been collected from states like Haryana, Punjab and western Uttar Pradesh. Similar trend was also reported from Greece, where 95% landraces were replaced by high yielding varieties (Rufo *et al.*, 2019). Landraces are widely used by international organization like CIMMYT and ICARDA for increasing the allelic diversity in wheat. During the past two decades advances in research techniques such as molecular techniques, geographical information system and information and communication technologies have given new insights into the utilization of genetic diversity, which can further facilitate the widening of the genetic base of wheat.

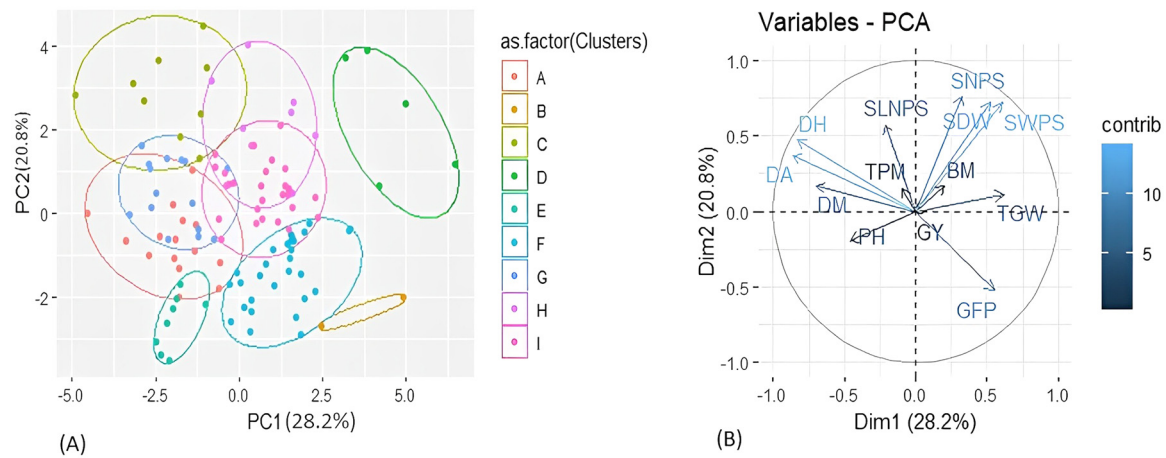
This cluster analysis indicates a considerable amount of genetic diversity present in the studied germplasm and can be used for classifying the genotypes. The cluster analysis confiscates genotypes into clusters which exhibit intra-cluster homogeneity and

inter-cluster heterogeneity. Cluster 'C' having nine genotypes mainly comprised of released varieties with higher values for yield, Dotlačil *et al.*, (2000) also reported one cluster was represented by checks, while studying the European winter wheat landraces. It is evident from the results that Cluster D is the most diverse cluster than others and comprised of wheat landraces collected from Uttarakhand. In Uttarakhand, wheat is cultivated under diverse agroecological situation (valley, mid hills, higher hills, rainfed, irrigated, early sown, timely sown etc.) and generally have longer vegetative growth period, relatively shorter grain filling period, and tall (Gupta and Kant, 2012). This is in conformity with finding of present work that cluster D is characterized by higher value for days to heading, PH and minimum period for grain filling. Among the two landraces in cluster B, Kapargaon is *T. sphaerococcum* (Indian Dwarf wheat). The *T. sphaerococcum* species is distinguished from other wheat species by dwarf PH,

**Table 3.** Cluster mean and standard deviation for different agro-morphological traits

Trait	Cluster								
	A(15)	B(2)	C(9)	D(5)	E(8)	F(30)	G(15)	H(9)	I(27)
DH (Days)	92 ± 3.3	98 ± 7.8	101 ± 2.2	114 ± 6.6**	87 ± 3.4*	96 ± 3.9	95 ± 2.8	103 ± 3.4	101 ± 2.7
DA(Days)	101 ± 2.9	107 ± 8.1	107 ± 4.9	121 ± 5**	98 ± 2.6*	104 ± 3.9	103 ± 3	112 ± 3.2	108 ± 3.2
DM (Days)	141 ± 2.3	140 ± 6.7	144 ± 2.9	153 ± 3.4**	140 ± 3.3	143 ± 2.7	139 ± 3.6*	145 ± 1.9	144 ± 2.4
PH (cm)	102.9 ± 7.7	96.5 ± 3.9*	98.9 ± 3.9	115.7 ± 4.7**	100.0 ± 9.2	113.4 ± 4.7	103.0 ± 6.9	112.4 ± 5.5	115.3 ± 7.8
SL (cm)	10.0 ± 1	6.52 ± 1.5*	10.3 ± 0.6	10.3 ± 0.4	10.5 ± 1.5	9.0 ± 0.8	10.9 ± 1.1**	9.5 ± 1.0	9.6 ± 0.9
SIPS (no.)	18.6 ± 1.1	19.0 ± 1.4	21 ± 0.8**	19.7 ± 1.0	17 ± 0.7*	18.3 ± 1	18.8 ± 1.1	19.6 ± 1	18.7 ± 1.1
SNS(no.)	46.3 ± 7.1	36.6 ± 4.7	51 ± 5.1**	37.2 ± 1.2	35.4 ± 5	35 ± 4.1*	43.8 ± 4.4	42.0 ± 4.6	41.8 ± 3.8
SWS (g)	2.0 ± 0.19	1.02 ± 0.02*	2.2 ± 0.29**	1.3 ± 0.4	1.7 ± 0.19	1.4 ± 0.16	1.9 ± 0.23	1.6 ± 0.18	1.8 ± 0.24
GFP(Days)	48 ± 2.6	42 ± 1.1	44 ± 1.5	39 ± 5.7*	53 ± 4**	48 ± 4	44 ± 4.4	42 ± 2.7	44 ± 3
TGW (g)	42.1 ± 3.5	24.99 ± 5.5*	39.9 ± 3.4	32.6 ± 8.7	45.6 ± 5.7**	37.0 ± 4.2	40.4 ± 1.9	36.5 ± 4.2	37.8 ± 3.0
GY (g)	281.2 ± 36.7	135.6 ± 24.6*	386.7 ± 55.3**	198.4 ± 49.2	288.3 ± 42.5	228.8 ± 32.5	341.1 ± 52	331.8 ± 79.6	239.1 ± 42.6
BM (kg/plot)	1.1 ± 0.11	0.63 ± 0.1*	1.2 ± 0.11	1.0 ± 0.07	1.1 ± 0.16	1.1 ± 0.11	1.2 ± 0.14	1.4 ± 0.2**	1.1 ± 0.14
TN (no.)	70.6 ± 6.3	64 ± 2.1*	80.3 ± 11.9	76.3 ± 9.9	79.9 ± 6.9	75.4 ± 10.2	83.9 ± 8.3	94 ± 6.8**	77.2 ± 7.8
SDW (g)	2.57 ± 0.3	1.44 ± 0.09*	2.94 ± 0.3**	1.81 ± 0.5	2.11 ± 0.2	1.91 ± 0.3	2.60 ± 0.3	2.18 ± 0.2	2.28 ± 0.3
CV %	8.34	11.2	7.55	12.3	8.87	8.69	8.51	9.15	8.63

Note: \*, \*\*lowest and highest means respectively; Value in parenthesis indicates the number of genotypes in that cluster.



**Figure 1.** (A) Scatter plot of 120 genotypes using PC1 and PC2 score. (B) Variable correlation plot.

small ear and semi-spherical wheat. In the present study also cluster B is characterized by dwarf PH and small ears. The archaeological remains recovered from historical site suggest that dwarf wheat is cultivated in Indian sub-continent during the period 2400-1750BC (Mittre, 1974). *T. sphaerococcum* is no longer

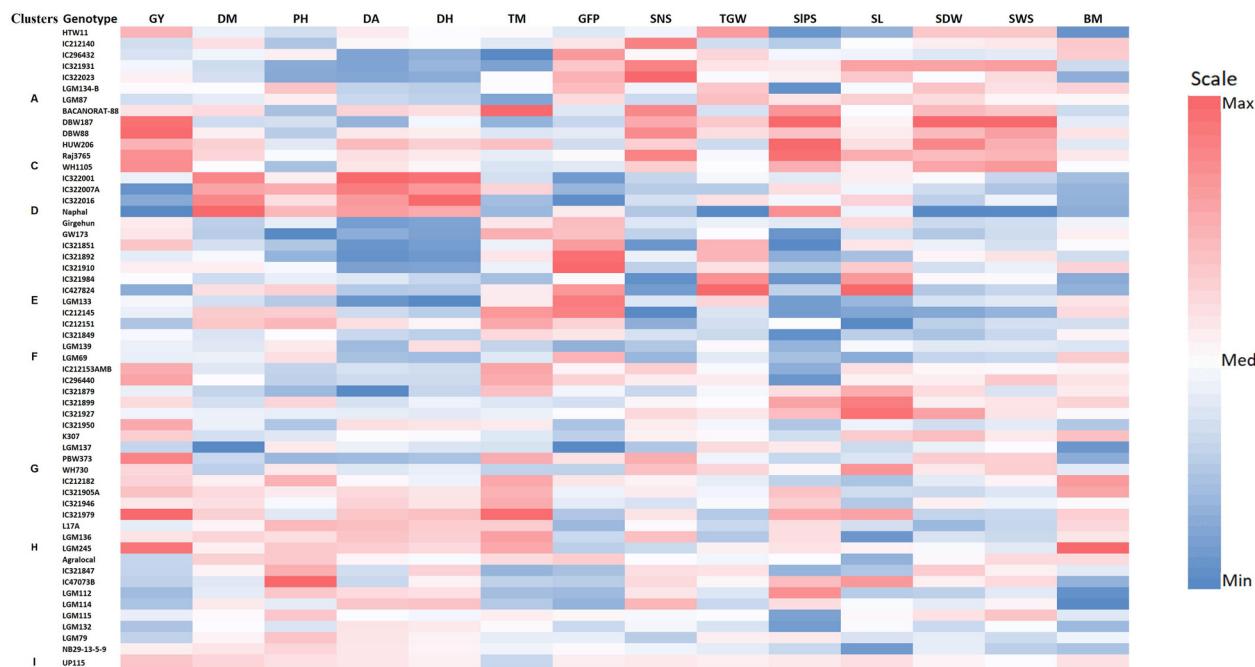
being cultivated in their native habitat but still being cultivated at 3–4 locations in Karnataka (Mori *et al.*, 2013). It is evident from the results that Cluster E and G comprised of early maturing landraces collected from M.P. and Gujarat with higher value for grain filling period. Mohan *et al.*, (2001) reported that the late

**Table 4.** Superior wheat genotypes identified for different agro-morphological traits

Traits	Criteria	Superior genotypes
DH	≤87 days	Landraces - LGM133(E), IC321892(E), IC321851(E), IC321910(E), Girgehun(E), RV - GW173 (E)
PH	≤92 cm	Landraces - IC321931(A), IC322023 (A) RV- GW173 (E)
GFP	≤40 days	Landraces - IC322016 (D), IC322001 (D), IC322007A (D); LGM114 (I), LGM112 (I), IC321847 (I), IC47073B (I); LGM137(G), IC321950 (G); L17A(H), IC321979 (H), LGM139(F)
SIPS	≥21	RV - DBW187(C), HUW206(C), Raj3765 (C); Landraces - LGM112 (I); Naphal (D)
SNS	≥50	Landraces - IC322023(A), IC212140(A), IC321931 (A); LGM136 (H); LGM114 (I) RV -Raj3765 (C), DBW88 (C), DBW187 (C); Exotic - BACANORAT-88, (C); PBW373 (G);
SWS	≥2.2 g	RV - DBW187 (C), WH1105 (C), DBW88 (C), HUW206 (C); Exotic - BACANORAT-88 (C); Landraces - IC321931 (A); LGM115 (I)
GY	≥375 g	Landraces - IC296440 (G), IC321979 (H), LGM245 (H), IC321856 (G) RV - WH1105 (C), HI977 (C), Raj3765 (C), PBW373 (G), DBW88(C), HD2967(C), DBW187(C)
SDW	≥3 g	RV - DBW187(C), HUW206(C), WH1105(C), Raj3765 (C); Landraces - IC321927 (G); IC321931 (A) GS - HTW11(A)
SL	≥12 cm	Landraces - IC321927 (G), IC321899 (G); IC427824 (E); GS - WH730 (G)
TN	≥93	Landraces - IC321979(H), LGM136(H), LGM245(H), IC212182(H), IC321946, IC321885 (H); IC212145(F), IC212151 (F); IC212153AMB (G), IC296440 (G) RV - PBW373 (G) Exotic - BACANORAT-88 (C)
BM	≥1.5 kg	Landraces - LGM245(H), IC212182(H), IC321905A (H); IC212140(A), IC296432 (A), LGM69 (F); RV - K307 (G), IC296491 (C)
TGW	≥45 g	Landraces - IC427824(E), IC321984(E), IC321892(E), IC321851 (E); IC321877 (F), LGM87 (A), LGM134-B (A) GS - HTW11(A); RV - DBW187 (C)

Note: Value in parenthesis indicates the cluster number.

RV, Released variety; GS, Registered Genetic Stock.



**Figure 2.** Heat map of superior genotypes identified for various traits.

sown but early maturing varieties are in great demand in central zone. The central zone encompasses MP and Gujarat states, and is considered as hot spot locations for heat stress in India (Roy and Poddar, 2022). The agro-ecological zone with recurrence stress factors tends to exhibit more homogenous genotypes and less degree of variation. Thus, the less value for intra cluster distance in cluster E and G may be due to recurrence of heat stress in M.P. and Gujarat during grain filling stage is common phenomenon. The wheat landraces evolved in hill ecosystem (cluster D) is considered too diverse from plain. The distance is so large that they formed separate diverse cluster. All modification occurring in crop plant during the course of evolution is dependent on selection pressure (Hammer, 2003). The selection pressure practiced in hills is based on maturity and height while farmers in plain selected the genotypes having more grain filling duration coupled with early maturity.

The cluster H and I contained landraces mainly collected from Madhya Pradesh, Gujarat, Rajasthan, Andhra Pradesh and North Eastern region of India. Cluster analysis did not group the landraces of same origin but grouped landraces together with greater agro-morphological similarity (Gebremariam *et al.*, 2022). Cluster analysis grouped different landraces together irrespective of origin indicating that geographical isolation may not be only factor causing diversity. The other factors responsible for such diversity are free flow of germplasm from one place to another (Verma and Mehta, 1976) genetic architecture of genotypes and climatic variability (Sharma and Hore, 1997).

The cumulative variance of 49.18% by first two principal component axes demonstrates that germplasm from each area has their own divergence and grouped separately. The independent grouping of all the accessions was due to variation among genotypes for different traits. Scatter plots separated different clusters in the present study and allowed natural clustering of wheat germplasm. Accordingly, the PCA and cluster analysis can be properly used for the clustering of wheat germplasm (Kraic *et al.*, 2009).

Thus, PCA-based cluster analyses indicate the difference more precisely among wheat germplasm rather than cluster analyses not based on PCA. The largest contribution to the total variation at each axis of differentiation is reflected by PCA. When data from all the traits were subjected to PCA, out of 10 only four principal component axes (PCA) had an Eigen value more than 1 and cumulatively explained 70.01% of the entire variability among genotypes, which gives a clear idea of structure underlying the variables analysed. According to Chahal and Gosal (2002), traits with the largest absolute value closer to unity within the first PC influence the clustering more than others that are closer to zero. Our results are in conformity with results reported by Gebremariam *et al.*, 2022, Poudel *et al.*, 2019 that first four PCs contributed nearly 72–76% of the total variation and had Eigen value more than unity among the wheat genotypes. The positive and negative correlation has been shown by positive and negative loading between the genotypes and traits. The trait which loads high value contributed more to the diversity. The significant positive correlation of PC1 with DH, DA, DM and PH was also reported by (Poudel *et al.*, 2017). The DH, DA, DM, SL, SIPS, SNS, SWS and SDW were the major traits of PC2 axis, hence, contributed more to diversity.

PCA method has been used by Wani *et al.* (2018) for selection purposes in terms of different traits for example grain yield, quality parameters etc. Based on their characteristics using multivariate statistical tools could help in selecting them based on graphical plots i.e. 2-D and 3-D diagrams. In the present studies, a principal component scatter plot of 120 wheat genotypes depicts that the genotypes which are close together are perceived as being similar when rated on the fourteen variables (Fig. 1(A)) and genotypes that are further at a distance are more diverse as explained by (Hussain *et al.*, 2014). When the vector value of the traits was superimposed on PC plot (Fig. 1(B)) DH & DA, SDW & SWS have a close angle between them representing a positive correlation. Trait DH and DA was opposite to GFP showing a negative

correlation. The angles are clear enough to draw a diagram of the correlation between the various indexes (Hussain *et al.*, 2014). From Fig. 1(B), it can be depicted that traits DH, DA, DM and PH contributed more to PC1 and trait GFP to PC2. The information obtained from this study can be used to plan crosses and maximized the use of genetic diversity and expression of traits as explained by Ahmad *et al.* (2013).

From the present study, a number of superior genotypes were identified for different traits. There were number of wheat landraces, which were early in DH, these landraces could be considered as adapted strategy for development of early maturing, climate resilient wheat varieties. In most cases, the superior genotypes for GFP, BM, TN and TGW were landraces. Alipour and Abdi (2020) also reported that landraces possess higher variation for BM and TGW. In comparison to released varieties, landraces were having more PH. Tall type indigenous wheat was also reported by Adhikari *et al.* (2022). GFP is an important parameter for temperature and drought tolerance and number of superior landraces of wheat were identified for less GFP. DBW187- a PAN India mega wheat for India released for three zones (NWPZ, NEPZ and CZ) has majority of desirable attributes. The Department of Agriculture and Cooperation got maximum breeder seed production indent of this variety during 2022–2023 (Singh *et al.*, 2022). Due to higher yield potential of released varieties, landraces are being replaced more rapidly from their area of cultivation. The released varieties exhibited superiority for grain yield over land landraces, while the landraces had higher values for TN, BM and TGW. Thus, for the improvement in TGW in released varieties, the hybridization between superior landraces for TGW from cluster E and released wheat varieties from cluster C could give desirable segregates.

The contribution of various characters toward the expression of divergence should be taken into account and analysis of the diversity of germplasm collections could be utilized for specific breeding goals. It is further suggested that multivariate methods can identify different parents, with the scope to generate transcendental segregates for future breeding strategies to improve wheat yields.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S1479262123000461>.

**Conflict of interest.** The authors declare that they have no conflict of interest.

## References

- Adhikari S, Kumari J, Jacob SR, Prasad P, Gangwar OP, Lata C, Thakur R, Singh AK, Bansal R, Kumar S, Bhardwaj SC and Kumar S (2022) Landraces-potential treasure for sustainable wheat improvement. *Genetic Resources and Crop Evolution* **69**, 499–523.
- Adilova SS, Qulmamatova DE, Baboev SK, Bozorov TA and Morgunov AI (2020) Multivariate cluster and principle component analyses of selected yield traits in Uzbek bread wheat cultivars. *American Journal of Plant Sciences* **11**, 903–912.
- Ahmad M, Shahzad A, Iqbal M, Asif M and Hirani AH (2013) Morphological and molecular genetic variation in wheat for salinity tolerance at germination and early seedling stage. *Australian Journal of Crop Science* **7**, 66–74.
- Ali Y, Khan MA, Hussain M, Atiq M and Ahmad JN (2019) An assessment of the genetic diversity in selected wheat lines using molecular markers and PCA-based cluster analysis. *Applied Ecology and Environmental Research* **17**, 931–950.
- Alipour H and Abdi H (2020) Interactive effects of vernalization and photoperiod loci on phenological traits and grain yield and differentiation of Iranian wheat landraces and cultivars. *Journal Plant Growth Regulation* **2**, 1.
- Al Lawati AH, Nadaf SK, Al Saady NA, Al Hinai SA, Almamari A, Al Adawi MH, Al Hinai RS and Al Maawali A (2021) Genetic diversity of Omani durum wheat (sub sp.) landraces. *Open Agriculture Journal* **15**, 21–32.
- Awan FK, Khurshid MY, Afzal O, Ahmed M and Chaudhry AN (2014) Agro-morphological evaluation of some exotic common bean (*Phaseolus vulgaris* L.) genotypes under rainfed conditions of Islamabad, Pakistan. *Pakistan Journal of Botany* **46**, 259–264.
- Beale EML (1969) Euclidian cluster analysis. Paper presented at the 37th session of International Statistical Institute, UK.
- Chahal G and Gosal SS (2002) *Principles and Procedures of Plant Breeding: Biotechnological and Conventional Approaches*. New Delhi: Narosa Publishing House.
- Degewione A and Alamerew S (2013) Genetic diversity in bread wheat (*Triticum aestivum* L.) genotypes. *Pakistan Journal of Biological Sciences* **16**, 1330–1335.
- Dempewolf H, Baute G, Anderson J, Kilian B, Smith C and Guarino L (2017) Past and future use of wild relatives in crop breeding. *Crop Science* **57**, 1070–1082.
- Dotlačil L, Hermuth J, Stehno Z and Manev M (2000) Diversity in European winter wheat landraces and obsolete cultivars. *Czech Journal of Genetics and Plant Breeding* **36**, 29–36.
- Gebremeriam K, Alamirew S and Gebreselassie W (2022) Evaluation of bread wheat (*Triticum aestivum* L.) germplasm at kafa zone, South-West Ethiopia. *Advances in Agriculture* **2022**, 1–7. <https://doi.org/10.1155/2022/1682961>
- Gupta HS and Kant L (2012) Wheat improvement in Northern hills of India. *Agricultural Research* **1**, 100–116.
- Hammer K (2003) A paradigm shift in the discipline of plant genetic resources. *Genetic Resources and Crop Evolution* **50**, 3–10.
- Hasan M, Hasibuzzaman ASM, Abdullah HM and Kallol MMH (2020) Genetic and genomic resources and their exploitation for unlocking genetic potential from the wild relatives. In Salgotra RK and Zargar SM (eds), *Rediscovery of Genetic and Genomic Resources for Future Food Security*. Singapore: Springer Nature, pp. 193–210. [https://doi.org/10.1007/978-981-15-0156-2\\_5](https://doi.org/10.1007/978-981-15-0156-2_5).
- Hussain SB, Wahid MA, Zubair M, Babar M and Wahid K (2014) Assessment of germplasm using multivariate analysis for grain yield and quality traits in spring wheat. *Pakistan Journal of Botany* **46**, 989–994.
- Jamali KD and Ali SA (2008) Yield and yield components with relation to plant height in semi-dwarf wheat. *Pakistan Journal of Botany* **40**, 1805–1808.
- Jaradat AA (2013) Wheat land races: a mini review. *Emirates Journal of Food and Agriculture* **25**, 20–29.
- Kraic F, Mocák J, Roháčik T and Sokolovičová J (2009) Chemometric characterization and classification of new wheat genotypes. *Nova Biotechnologica* **9**, 101–106.
- Maniee M, Kahrizi D and Mohammadi R (2009) Genetic variability of some morpho-physiological traits in Durum wheat (*Triticum turgidum* var. durum). *Journal of Applied Science* **9**, 1383–1387.
- Mecha B, Alamerew S, Assefa A, Assefa E and Dutamo D (2017) Genetic diversity based on multivariate analysis for yield and its contributing characters in bread wheat (*Triticum aestivum* L.) genotypes. *Agriculture Research & Technology: Open Access Journal* **8**, 1–10.
- Mittre V (1974) Palaeobotanical evidence in India. In Hutchinson J (ed.), *Evolutionary Studies in World Crops*. Cambridge: Cambridge University Press, pp. 3–30.
- Mohammadi SA and Prasanna BM (2003) Analysis of genetic diversity in crop plants- Salient statistical tools and considerations. *Crop Science* **43**, 1234–1248.
- Mohan D, Nagarajan S, Singh RV and Shoran J (2001) Is the national wheat breeding programme demand-driven?—An analysis. *Current Science* **81**, 749–753.
- Mori N, Ohta S, Chiba H, Takagi T, Niimi Y, Shinde V and Osada T (2013) Rediscovery of Indian dwarf wheat (*Triticum aestivum* L. ssp. *sphaerococcum* (perc.) mk.) an ancient crop of the Indian subcontinent. *Genetic Resources and Crop Evolution* **60**, 1771–1775.
- Mujaju C and Chakuya E (2008) Morphological variation of sorghum landrace accessions on-farm in semi-arid areas of Zimbabwe. *International Journal of Botany* **4**, 376–382.
- Panwar NS, Bhatt KC, Dhariwal OP, Pandey A and Jacob S (2014) Spatial distribution of trait-specific diversity in Indian wheat collections. *Indian Journal of Plant Genetic Resource* **27**, 280–286.



- Peng JH, Sun D and Nevo E** (2011) Domestication evolution, genetics and genomics in wheat. *Molecular Breeding* **28**, 281–301.
- Poudel A, Thapa DB and Sapkota M** (2017) Assessment of genetic diversity of bread wheat (*Triticum aestivum* L.) genotypes through cluster and principal component analysis. *International Journal of Experimental Research and Review* **11**, 1–9.
- Poudel R, Finnie S and Rose DJ** (2019) Effects of wheat kernel germination time and drying temperature on compositional and end-use properties of the resulting whole wheat flour. *Journal of Cereal Science* **86**, 33–40.
- Priya B, Diyali S, Mukherjee S and Srinivasarao M** (2015) Genetic diversity based on cluster and principal component analysis in wheat and triticale genotypes. *Research on Crops* **16**, 712–718.
- R Core Team** (2020) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R foundation for statistical computing.
- Reitz LP and Salmon SC** (1968) Origin, history, and use of Norin10 wheat. *Crop Science* **8**, 686–689.
- Roy S and Poddar S** (2022) Effect of terminal heat stress in wheat: an overview. *Agriculture and Environment* **3**, 15–20.
- Rufo R, Alvaro F, Royo C and Soriano JM** (2019) From landraces to improved cultivars: assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers. *PLoS One* **14**, e0219867.
- Sharma BD and Hore DK** (1997) Genetic divergence in Himalayan rice. *Indian Journal of Plant Genetic Resources* **10**, 181–185.
- Shekhawat US, Vijay P and Singhania DL** (2001) Genetic divergence in barley (*Hordeum vulgare* L.). *Indian Journal of Agricultural Research* **35**, 121–123.
- Simmonds NW** (1979) Principles of crop improvement. Longman, Newspaper.
- Singh G, Tiwari R, Tyagi BS, Gupta A, Kumar S, Khan H, Gupta V, Sharma AK, Mishra CN, Kumar V, Singh C, Mamrutha HM, Kamble UR, Verma A and Singh GP** (2022) Progress Report of AICRP on Wheat and Barley 2021-22, Crop Improvement. ICAR-Indian Institute of Wheat and Barley Research, Karnal, Haryana, India, p.219.
- Sparks DN** (1973) Euclidian cluster analysis algorithm. *Journal of Applied Statistics* **22**, 126–130.
- Tandon JP** (2000) Wheat breeding in India during the twentieth century. *Crop Improvement* **27**, 1–18.
- Verma VS and Mehta RK** (1976) Genetic divergence in lucerne. *Maharashtra Agriculture University* **1**, 23–28.
- Villa TCC, Maxted N, Scholten M and Ford-Lloyd B** (2005) Defining and identifying crop landraces. *Plant Genetic Resources: Characterization and Utilization* **3**, 373–384.
- Wani SH, Sheikh FA, Najeeb S, Iqbal AM, Kordrostami M, Parray GA and Jeberson MS** (2018) Genetic variability study in bread wheat (*Triticum aestivum* L.) under temperate conditions. *Current Agriculture Research Journal* **6**, 268–277.
- Yadav SK, Singh AK and Malik R** (2015) Genetic diversity analysis based on morphological traits and microsatellite markers in barley (*Hordeum vulgare*). *Indian Journal of Agricultural Science* **85**, 37–44.
- Zhong GY and Qualset CO** (1995) Quantitative genetic diversity and conservation strategies for an allogamous annual species, *Dasyphyrum villosum* (L.) Candargy (Poaceae). *Theoretical and Applied Genetics* **91**, 1064–1073.