Original Research Article

Deciphering Genetic Diversity in Advanced Wheat Lines (*Triticum aestivum* L.) for Yield and Other Yield Contributing Traits across the Locations

Abstract

This study aimed to explore the genetic diversity among 20 wheat genotypes using principal component analysis (PCA) and cluster analysis, providing a comprehensive understanding of the present genetic variations. In this study, the agro-morphological characteristics of the genotypes were evaluated across two distinct locations using a randomized complete block design (RCBD) with three replications. Three different clusters were formed, with the maximum number of genotypes occurring in the second cluster. A lower pvalue indicates that the clusters are statistically significant, suggesting that the observed diversity is not random. High estimates of cophenetic distance (.79) specify a high genetic distance between clusters, indicating that there is diverse genetic material under study. The maximum genetic distance observed was 144.9 between genotypes 5 and 11. These findings suggest that these two genotypes are the most genetically diverse among all the studied genotypes and can be used as parents to develop high genetic variation in the studied traits. The PCA results yielded 18 principal components, with the first seven components accounting for approximately 83% of the total variance, indicating significant genetic diversity. The scree plot affirmed the robustness of the study's results by suggesting that the first eight principal components accounted for a substantial portion of the variance. These genetic distances between genotypes provide valuable insights for crop improvement, aiding in the development of more effective breeding strategies. Therefore, PCA offers valuable insights into the genetic diversity of wheat lines, which could be beneficial for future breeding programs.

Key words: Wheat, Genetic diversity, PCA, Cluster andcophenetic distance

1. INTRODUCTION

Wheat (*Triticum aestivum* L.) is a globally significant crop that contributes substantially to food security worldwide, particularly in India. It thrives in warm temperature regions, with temperatures ranging from 21 to 24°C. As the most widely cultivated crop in the world, wheat is cultivated on 220 million hectares annually, yielding a total production of 788.5 million tones (FAO, 2023). World wheat production has been lifted by 1.4 million tonnes and now stands at 788.5 million tonnes in 2023

Wheat plants serve as a staple food in many cultures and provide approximately 20% of the total dietary calories and proteins worldwide (Leiva, 2022). Its economic importance is immense, as it servesas a staple food for millions of people. Genetic diversity is of paramount importance in crop improvement. This approach provides a pool of variation from which plant breeders can select traits that enhance crop yield, disease resistance, and environmental adaptability. However, the use of breeding lines with narrow genetic bases can lead to significant losses (Kumar et al., 2022). Cultivating germplasm with a narrow genetic base carries a risk of genetic vulnerability. Mutations in pest populations or changes in environmental conditions can introduce stresses that the cultivar cannot cope with, potentially leading to severe crop losses. Therefore, maintaining and utilizing a broad genetic base in crop breeding is crucial for sustainable agriculture and food security.

Different algorithmic methods, such as multidimensional scaling, clustering, principal component analysis, and principal coordinate analysis, are currently employed in the assessment of genetic diversity(Rohlf, 1988:Thompson et al., 1998; Melchinger, 1993; Brown-Guedira et al., 2000).Principal component analysis (PCA), a dimensionality reduction technique introduced by Karl Pearson in 1901, identifies the largest variations in the data. Cluster analysis, a method of grouping objects based on similarity, is often used in conjunction with PCA. These techniques play a pivotal role in crop improvement strategies, including in wheat, by determining germplasm variability. These methods assist in identifying lines for desirable traits, segregating progenies with maximum genetic variability for further selection, and introducing desirable genes from diverse germplasms into the available genetic base. These techniques ensure continued improvement in plant selection programs and have been used to study the genetic diversity and relationships of wheat genotypes, which is crucial for planning crosses, assigning lines to specific heterotic groups, and precisely identifying plant varietal protection (Govindaraj et al., 2015; Khodadadi et al., 2011; Szczepanik et al., 2021; Bhandari et al., 2007).

The significance of genetic diversity in wheat is welldocumented (Gruet et al., 2022;Khan et al., 2015, Yadav et al., 2014).The genetic variability among plants dictates their potential for improving efficiency and consequently, their suitability for breeding programs, which could ultimately result in increased food production(Zedfar and Golabadi, 2013).This diversity paves the way for plant breeders to cultivate improved varieties with desirable traits such as high yield potential, large grains, and resistance to biotic and abiotic stresses (Atsbeha et al., 2023). Therefore, the exploration and utilization of genetic diversity in wheat genetic resources are vital for sustainable production. Genetic diversity is a cornerstone of crop breeding because it augments yield potential by integrating desirable traits from diverse parents (Joshiet al., 2023; Chaudhary et al., 2022; Elahi et al., 2021). Moreover, a diverse genetic base endows crops with enhanced resilience against climatic changes, thereby ensuring sustainable agricultural production amidst environmental uncertainties. This resilience has gained particular importance as climate change poses an increasing threat to agricultural production (Joshi et al., 2023). In this study, we employed cluster analysis and principal component analysis (PCA) to probe the genetic diversity of advanced wheat lines, focusing on their agronomic traits. Our research aimed to identify elite wheat breeding lines, thereby enriching the understanding of wheat genetic diversity and paving the way for future advancements in wheat breeding.

EXPERIMENTAL DETAILS

With a completely randomized block design (RBD), twenty advanced wheat genotypes were assessed with three replications across six field trials during two cropping seasons, Rabi 2020-21 and 2021-22. These trials took place at the N.E. Borlaug Crop Research Centre (NEBCRC), G.B. Pant University of Agriculture and Technology, Pantnagar, District U.S. Nagar, and another location, the Agriculture Research Center Majhera, Nainital, Uttarakhand.

Comment [KS1]: The paragraph provides a detailed overview of the experimental setup, including the design, locations, crops studied, traits measured, and analytical methods used. However, it would be beneficial to clarify certain aspects such as the specific wheat genotypes used, the exact procedures followed for data collection and analysis, and any potential confounding factors that were controlled for during the experiments.

The plants of each genotype were planted in a 4 row plot, each of which was 4 meters long, with rows spaced 20 cm apart. All recommended wheat cultivation practices were followed for a healthy crop. Eight quantitative traits were observed from five randomly selected plants in each entry. These traits included flag leaf length (FLL), flag leaf width (FLW), exposed peduncle length (EPL), total peduncle length (TPL), spike length (SL), awn length (AL), plant height (PH), and spikelets per spike (SPS). On a plot basis, observations were recorded for germination percentage (GP), seedling vigor (SV), days to anthesis (DA), days to heading (DH), and days to maturity (DM). The total number of productive tillers (TPM)counted from area of one meter in row. 1000-grain weight (TGW) and yield per plot (YPP) were recorded from clean harvest. Brown and yellow rusts were observed as described byPeterson et al. (1948).

Cluster analysis was performed using the unweighted pair group method with arithmetic mean (UPGMA), as proposed by Sokal and Michener (1958), to assemble a phylogenetic tree from a distance matrix. Genetic distances between the genotypes were calculated using the Euclidean method (Gower, 1985).

Principal component analysis (PCA), a statistical tool commonly used in plant breeding to identify trends in multidimensional data (Pearson, 1901 and Hotelling, 1933), was employed. This approach aids in studying the morphological characteristics of germplasm, assessing population differences and breeding potential and reducing data redundancy (Khodadadi et al. 2011; Sewell 2008). PCA and cluster analysis, as well as visualization, were performed using R script (Mojena, 1977; Kassambara, 2020; Le et al. 2008; Husson et al. 2017).

3. Results and discussion

3.1. Cluster analysis

The results of the cluster dendrogram for genotypes with P values (AU is an approximately unbiased P value computed by multiscale bootstrap resampling, and BP is the bootstrap probability value computed by normal bootstrap resampling) are shown in **Fig.1**. The 20 genotypes were grouped into clusters based on their genetic similarity for the agronomic characteristics under study. The results showed that the whole genotypes were grouped into two groups and that genotype 11 was an outlier. The cophenetic distance was 0.79, which represents the distance between two clusters, indicating that the

Comment [KS2]: References to previous research studies strengthens the validity of the findings and situates the current study within the broader context of wheat genetics and breeding. However, it would be beneficial to briefly summarize the key findings of these previous studies and discuss how they relate to the findings of the current study

Comment [KS3]: The language used is clear and technical terms are appropriately defined. However, it would be helpful to ensure consistency in the formatting of genotype labels (e.g., consistently referring to genotypes as "G1," "G2," etc.) throughout the text as somewhere its written G1 somewhere its genotype 1 studied genotypes are highly diverse in terms of their traits. p values of 7 and 18 for clusters two and three, respectively, inferred that clusters are not formed by chance and that there is genuine diversity in the variables used between them.

The first branch of the dendrogram included genotypes G11, G7, G5, G13 and G4, which formed cluster one with genotypes G13 and G4. These genotypes are closely related, as indicated by the short branch lengths and second branch again grouping into two subgroups, which form two additional clusters. Cluster two contained G1, G19, G17, G6, G2, G14 and G3 and had the highest number of genotypes in this cluster. Genotypes 12 and 16, which form the third cluster indicating 20 advanced genotypes, are genetically diverse. Other genotypes form individual pairs or small clusters, indicating specific close relationships between those genotypes. The results of the dendrogram inferred that there was a great amount of genetic diversity among the 20 genotypes. The three main clusters (cluster one, cluster two and cluster three) represent groups of genotypes that are genetically similar within the group but distinct from each other. This genetic diversity could be leveraged in breeding programs to introduce new traits or improve existing traits. The cluster genotypes, including G13or G4, can be used as parents to cross genotypes from cluster two (G1, G19, G17, G6, G2, G14 and G3) or from the third cluster (G12 and G16) to maximize genetic diversity and potentially introduce beneficial traits that are present in one cluster but not the other. Alternatively, genotypes within the same cluster could be crossed to reinforce specific traits that are common within that cluster.

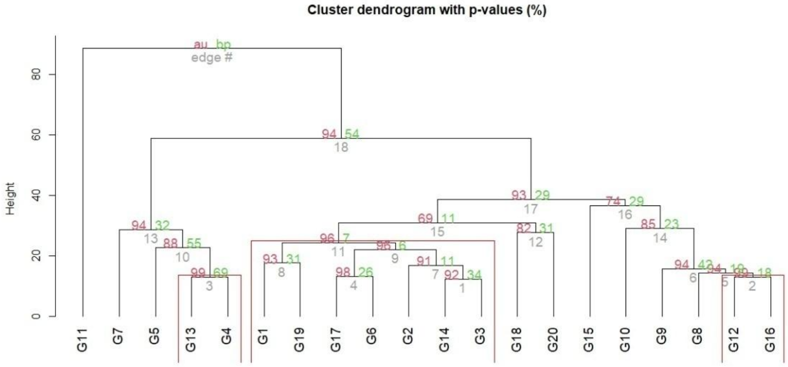

Fig 1. Hierarchical clustering dendrogram illustrating the genetic diversity among 20 wheat genotypes

.

The maximum genetic distance observed was 144.9 between genotypes 5 and 11 **(Table 1).**These findings suggested that these two genotypes are the most genetically diverse among all the studied genotypes and can be used as parents to explain large amounts of genetic variation in the studied traits. On the other hand, the minimum non-zero genetic distance was 12.32 between genotypes 3 and 14, indicating that these two groups are the most genetically similar among all the studied genotypes. The second largest genetic distance was 128.1 for genotype pairs 4 and 11, followed by 113.7, 110.1 and 102.4 for genotype pairs 11&7, 10&5 and 13&10, respectively. A second lowest value of genetic distance was recorded for genotype pairs 13 and 4, followed by 17 and 6.3, 214 and 6, 12 and 18. Similar results were previously reported by Chaudhary et al. (2022),Santosh et al. (2019),Jaiswal et al. (2010).Itis recommended that genotypes that are genetically diverse (i.e., have a large genetic distance) be crossed to maximize genetic diversity in the offspring. Conversely, if the goal is to maintain certain traits, breeders might choose to cross groups that are genetically similar (i.e., have a small genetic distance). By utilizing this information for crop improvement, genetic distance information can provide valuable insights into the genetic diversity and relationships among different genotypes and help to make informed decisions and develop more effective breeding strategies.

Studies conducted previously by Khalid et al. (2022), Braved et al. (2022), Vuset al. (2020)and Shimelis et al. (2018)also implied similar results**,** suggestingthat the variation in genotypes across clusters resulted from the minor impact and cumulative influence of several characteristics and had the potential for hybridization programs to design crosses for the manifestation of heterosis and for improving quality traits.The utilization of cluster analysis in assessing the genetic resources have proven instrumental in identifying valuable starting material for priority breeding areas.The detailed breakdown of clusters in the previous studies highlights the relevance of cluster analysis in delineating distinct groups with specific trait combinations. This knowledge can be applied to wheat breeding, allowing researchers to prioritize and select cultivars with optimal trait combinations for targeted environments.

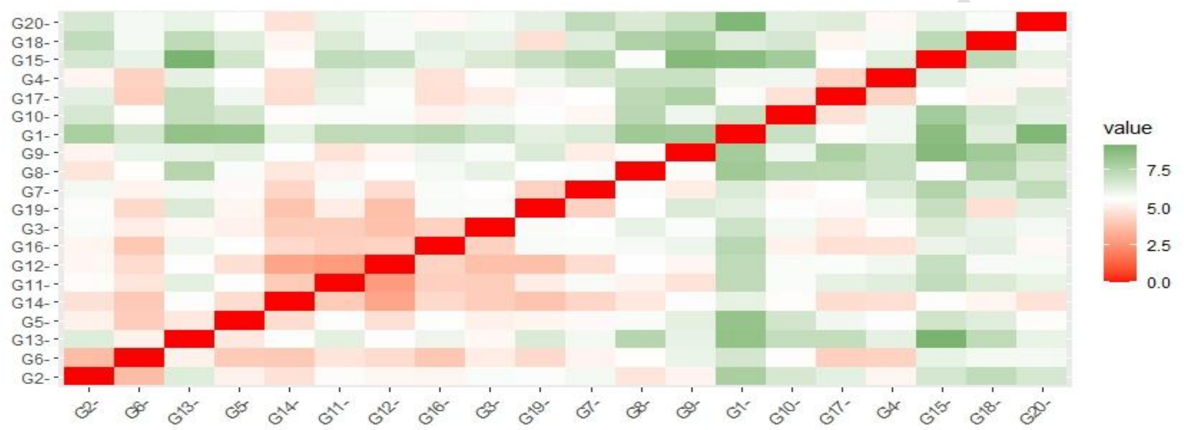

Fig 2. Heatmap representing the genetic distance of 20 advanced lines of wheat

Comment [KS4]: Describing what Heatmap implies here will provide the in depth understanding. And how this map provide significance on your study. Is it like how the genotype is related with each other or correlation among any typical trait or? something like this to corelate your image with text.

Table 1. Genetic distance matrix of 20 wheat genotypes showing genetic diversity for agromorphological characteristics

	G1	**G2**	**G3**	**G4**	**G5**	**G6**	**G7**	**G8**	**G9**	**G10**	**G11**	**G12**	**G13**	**G14**	**G15**	**G16**	**G17**	**G18**	**G19**	**41.3**
G1	0.0	26.5	20.1	51.2	71.4	25.8	39.5	34.6	31.1	52.0	85.1	27.9	60.1	22.8	50.0	38.6	32.5	32.8	17.6	41.3
G2		0.0	13.4	32.1	50.5	29.6	24.2	46.2	47.3	70.8	103.3	40.4	41.1	20.4	51.6	50.5	31.3	31.2	32.5	44.5
G3			0.0	36.9	56.3	20.2	26.8	38.3	38.1	60.2	93.8	30.3	46.7	12.3	43.9	40.4	23.0	26.7	23.1	35.9
G4				0.0	24.7	48.8	21.9	70.8	71.3	92.4	128.1	64.4	13.0	38.5	64.0	72.4	42.3	37.7	55.6	52.1
G5					0.0	64.7	34.5	86.8	89.6	110.1	144.9	81.4	20.7	56.1	74.3	88.7	56.1	55.5	73.8	68.2
G6						0.0	34.4	24.5	27.8	46.9	80.6	18.5	58.9	13.4	26.7	25.6	13.2	30.8	16.2	26.4
G7							0.0	55.6	57.3	78.5	113.7	49.7	29.5	24.7	50.0	58.3	27.9	31.3	41.1	43.6
G8								0.0	14.5	31.1	60.1	13.7	81.0	33.4	31.3	15.1	35.7	48.8	21.8	37.9
G9									0.0	25.6	59.1	14.6	81.1	34.9	40.4	18.3	39.3	47.6	19.6	37.2
G10										0.0	39.4	32.9	102.4	56.1	49.6	26.6	55.8	65.3	40.6	49.9
G11											0.0	64.7	138.3	90.7	78.0	57.0	90.2	100.0	74.9	85.2
G12												0.0	74.4	26.7	32.3	13.0	29.3	40.4	13.9	33.2
G13													0.0	48.6	74.3	82.9	52.4	45.9	65.1	61.8
G14														0.0	34.8	36.0	15.0	23.1	18.8	27.2
G15															0.0	30.0	26.3	47.0	36.7	30.9
G16																0.0	34.8	48.1	24.9	35.4

G17	0.0	28.2	25.6	26.5
G18		0.0	30.4	27.7
G19			0.0	29.6
G20				0.0

3.2 Principal component analysis

Comment [KS5]: Ensuring consistency in formatting (e.g., capitalization of "Principal Component Analysis") would improve readability.

Principal component analysis (PCA) revealed that a certain amount of diversity was present in the studied genotypes for the agromorphological traits, resulting in 18 principal components (**Table 2**). The first seven components captured a significant portion of the genetic diversity among the wheat genotypes for the studied traits. These seven components explain approximately 83% of the total variance in the data, suggesting that they account for the majority of the genetic diversity in the wheat genotypes studied. The first principal component explained 20.54% of the total variance, with an eigenvalue of 3.70, while the second principal component accounted for an additional 17.46% of the variance. As a result, the cumulative variance explained 37.99% of the variance by the addition of principal component three, thereby reaching half of the cumulative variance present in the genotypes. Similar results for PCA were also observed byAbdelghany et al. (2023), Adilova et al. (2020),Shivramakrishnanet al. (2016) and Poudel et al. (2017).

The contributions of the various variables to the top five principal components for the variation in the 20 advanced lines of wheat are listed in **Table 3**. In principal component (=Dimension) 1, the spike length had the highest contribution (13.44), while the germination percentage had the lowest (0.14), suggesting that spike length is a significant factor in differentiating the wheat lines along this dimension compared to germination percentage. After spike length, peduncle length (both exposed and total) was the major contributing variable in principal component 1. Days to maturity were the leading contributor, while spikelets per spike were the least contributing factor. Similarly, in PC3, PC4, and PC5the major contributing variables were brown rust, flag leaf length, and yield per plot respectively, suggesting that these variables are deciding factors for variation among genotypes and that targeting improvements in these characteristics could lead to improvements in the genotypes. The same analysis can be applied to the other dimensions. It is also worth noting that the variables contributing the most to each dimension are different. Earlier studies of principal components in wheat were performed by Shamuyarira et al. (2022); Sharma et al.(2021); Singh et al.(2020); Riaz et al. (2018)and Amin et al. (2014), who concluded that principal component analysisis an excellent tool for determining genetic diversity. This suggests that each dimension captures a different aspect of the variation in the data. Overall, these results provide valuable insights into the genetic diversity of wheat lines, which could be useful for future breeding programs. By understanding which traits contribute most to the variability, breeders can focus on these traits when selecting lines for cross-breeding. Variables such as spike length (0.50), total peduncle length (0.47), and exposed peduncle length (0.41) had high communalities, indicating that these variables were wellrepresented by the extracted factors.

The exploration of Principal Component Analysis (PCA) in wheat research has garnered considerable attention, with several studies delving into its application and implications**.**Piro et al. (2023), illustrated the potential application of rye chromatin introgression in wheat quality breeding, with the arabinoxylan content of wheat white flour, demonstrating the use of PCA in assessing quality traits in wheat breeding Kumar et al. (2021),foundvariability in agro-morphological and physiological traits by PCA, suggesting that identified genotypes can be used for hybridization and improved cultivar development. Saleh et al. (2021) emphasized the use of PCA in assessing genetic variations among wheat genotypes to enhance selection efficiency in breeding programs.Moreover, Ahmad et al. (2017) conducted PCA to examine the suitability of wheat varieties for cookie making quality, demonstrating the use of PCA in correlating physical and rheological parameters of wheat varieties.These studies collectively demonstrate the diverse applications of PCA in wheat research, including assessing wheat variety suitability for specific products, comparing wheat genotypes, and analyzing wheat flour refinement. The use of PCA in these studies highlights its effectiveness in providing valuable insights into the composition, properties, and genetic diversity of wheat, thereby contributing to advancements in wheat breeding and product development.

Table 2.Eigenvalues, proportions of variance and cumulative proportions of 20 wheat genotypes

Principal Components	Eigen value	Percentage of Variance	Cumulative Percentage of Variance
comp1	3.70	20.54	20.54
comp2	3.14	17.46	37.99
comp3	2.38	13.22	51.22
comp4	1.72	9.56	60.77
comp5	1.63	9.04	69.81
comp6	1.22	6.79	76.60
comp7	1.14	6.32	82.92
comp8	0.92	5.09	88.01
comp9	0.51	2.85	90.86
comp10	0.49	2.73	93.58
comp11	0.41	2.30	95.88
comp12	0.33	1.84	97.72
comp13	0.23	1.28	99.01
comp14	0.11	0.64	99.64
comp15	0.05	0.28	99.92
comp16	0.01	0.07	99.98
comp17	0.00	0.01	100.00
comp18	0.00	0.00	100.00

Table 3. Contribution of the top five principal components to the variation in 20 advanced lines of wheat

Variable	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
GP	0.14	7.80	14.19	4.89	0.43
SV	3.56	6.86	8.75	2.08	0.21
DA	7.26	14.18	5.06	0.72	2.80
DH	8.50	13.21	4.07	1.72	2.07
DM	3.21	16.31	1.34	1.28	0.12
FLL	6.99	0.00	1.22	21.71	0.28
FLW	9.63	0.94	2.61	12.29	2.76
EPL	11.08	8.51	0.16	2.84	0.84
TPL	12.60	7.56	3.33	0.80	3.69
SL	13.44	0.74	3.93	8.99	0.00
AL	0.14	3.52	0.35	6.67	5.03
PH	7.03	0.17	3.01	7.83	18.69
SPS	10.44	0.01	13.58	1.79	0.02
TPM	0.43	15.25	3.77	0.14	0.33
TGW	0.00	3.85	0.04	6.50	24.26
YPP	0.22	0.01	7.77	0.65	35.36
YR	4.57	0.48	8.05	18.20	1.88
BR	0.76	0.60	18.77	0.92	1.24

The scree plot (Figure 2) graphically represents the percentage of variance explained by different principal components. The results indicate that the first principal component explains 20% of the variance, while the second principal component accounts for slightly less than 20%. The scree plot suggests that considering the first eight principal components would capture a substantial amount of the variance in the study. These results are consistent with those ofAmbati et al. (2020), Mishra et al. (2015),Sarfrazet al. 2021.Therefore, the robustness of the study's results can be attributed to these eight principal components.

Comment [KS6]: The explanation of how the scree plot assists in determining the number of principal components to consider is clear and understandable. However, it would be beneficial to provide a brief explanation of how the scree plot is constructed and how it aids in determining the optimal number of principal components.

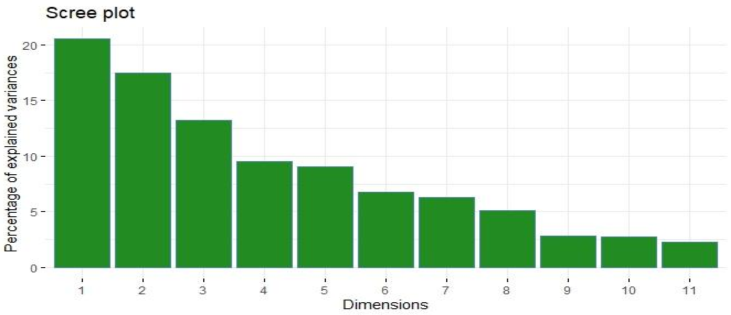

Fig. 3 shows a scree plot diagram built based on eleven principal components.

Dimension 1 explained 20.5% of the variance in genotypes for traits, while dimension 2 accounted for 17.6% of the variance (**Fig 4& Fig 5**). The biplot features multiple vectors, each representing all the characters under study. The direction and length of each vector indicate how each variable contributes to the two principal components. A longer vector indicates a variable that strongly influences the score of the individuals on the corresponding principal component. The results showed that day to anthesis, heading and maturity were positively correlated, and selecting one characteristic directly improved the other characteristics. Similar tiller per meter and germination percentage values are positively associated, suggesting that improving the germination percentage directly benefits the tiller per meter. However, the number of tillers per meter was negatively correlated with awn length and thousand-grain weight.The color scale indicates the contribution of each variable to the principal components, with darker colors representing greater contributions(**Fig.4)**.

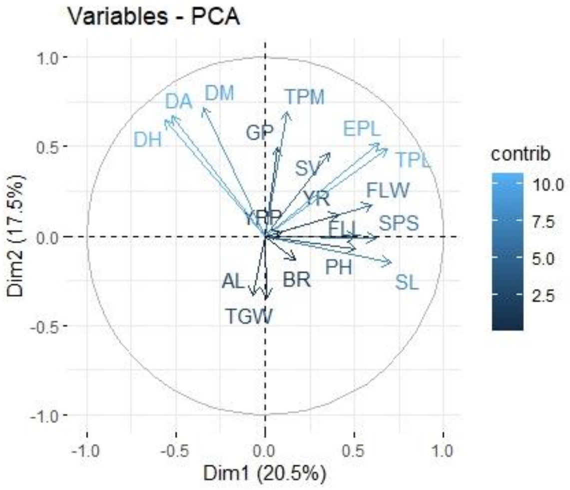

Fig 4. Principal component biplot for variables in 20 advanced genotypes of wheat

The PCA biplot for individual plots represents an individual, labeled from 1 to 20. The distribution of points across all four quadrants indicates variation among the individuals along these two principal components. The exact position of each point provides information about the individual's scores on the two components. For instance,Genotype 7 was located in the top right quadrant, in which higher values were observed for both dimensions (**Figure 5**). This finding suggested that genotype 7 had high scores for both Dim1 and Dim2. The contributions of genotypes 20, 5, 18 and 4 were negative, and likewise their contributions were lower. Genotypes 1, 2 and 10 positively contributed to dimension 1. This PCA plot can provide valuable insights into the underlying structure of the data, helping to identify patterns and relationships among individuals who might not be apparent from the raw data alone.

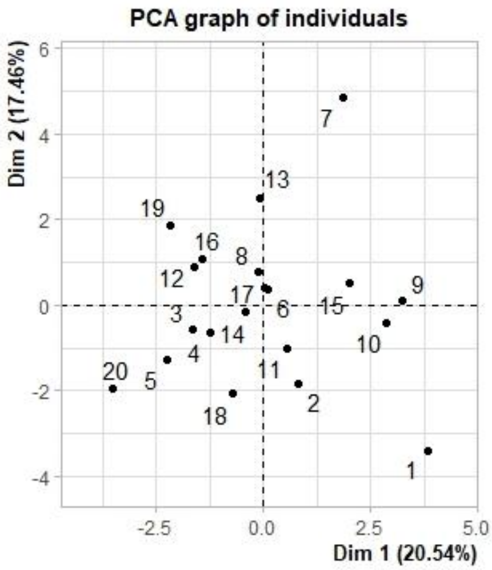

Fig 5. PCA biplot for individualsin 20 advanced genotypes of wheat

3.3 Rotated component matrix analysis

The rotated component matrix scores for Principal Component 1 (PC1) across the 20 wheat genotypes revealed a diverse set of characteristics**(Fig. 6).** The scores ranged from a high of 1.94 (genotype 1) to a low of -1.79 (genotype 20). These scores represent the correlation of each genotype with the first principal component, which is a linear combination of the original variables that captures the maximum variance in the data.

Comment [KS7]: The study effectively communicates the results of the rotated component matrix analysis and their significance for understanding genetic diversity among wheat genotypes.
With the integration of previous studies and further elaboration on the practical implications of the findings, the paragraph would enhances its contribution to the field of wheat genetics and breeding.

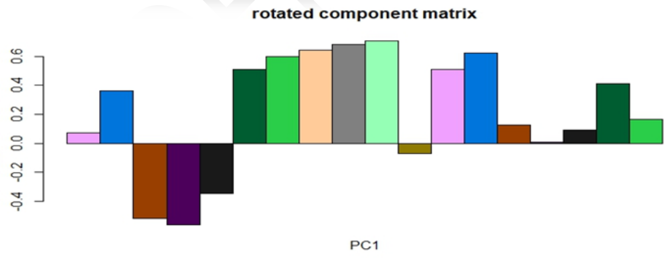

Fig.6. Rotated component metrics (PCs)1 of 20 wheat genotypes.

Positive scores on PC1 (e.g., genotypes 1, 2, 6, 7, 9, 10, 11, and 17) suggested that these genotypes shared certain traits or characteristics. In contrast, genotypes with negative scores on PC1 (e.g., genotypes 3, 4, 5, 12, 16, 18, 19, and 20) indicate different or opposing traits compared to those with positive scores. Similar results were observed by Sheela et al.(2020), Singhet al. (2020),Nachimuthu et al. (2014) and Kumar et al.(2014). These results highlight the genetic diversity among the studied wheat genotypes, which is crucial for crop improvement strategies. The variability in PC1 scores can be used to identify specific traits for further selection and to introgress desirable genes from diverse germplasms into the available genetic base.

CONCLUSION

Plant breeding relies on the principle of continuous selection of favorable alleles present in the available germplasm. This selection aims to improve the performance of cultivars or elite lines against major biotic or abiotic stresses in their environment. In this context, the study successfully classified 20 wheat genotypes into three clusters based on their genetic similarity, with a high cophenetic distance of 0.79 and low pvalues for the clusters suggesting high diversity for the studied traits. This diversity can be harnessed in breeding programs to introduce new traits or enhance existing traits and can be further elucidated through principal component analysis (PCA). The PCA revealed considerable diversity among the studied wheat genotypes for agro-morphological traits, with the first seven components accounting for approximately 83% of the total variance, indicating significant genetic diversity. The scree plot affirmed the robustness of the study's results by suggesting that the first eight principal components accounted for a substantial portion of the variance. Notably, Genotype 7, located in the top right quadrant, demonstrated high scores on both dimensions, indicating its significant contribution to genetic diversity, while genotypes 20, 5, 18, and 4 contributed less to genetic diversity, with negative values. Genotypes 1, 2, and 10 positively contributed to dimension 1. These genetic distances between genotypes provide valuable insights for crop improvement, aiding in the development of more effective breeding strategies. Therefore, PCA offers valuable insights into the genetic diversity of wheat lines, which could be beneficial for future breeding programs.

CONSENT (WHERE EVER APPLICABLE)

ETHICAL APPROVAL (WHERE EVER APPLICABLE)

Reference

Abdelghany M, Makhmer K, Zayed E, salama Y and Amer K. Genetic Variability, Principle Components and Cluster Analysis of Twenty-eight Egyptian Wheat Genotypes. *Scientific Journal of Agricultural Sciences*. 2023;*5*(1), 107-118. doi: 10.21608/sjas.2023.179573.1272

Adilova SSh, Qulmamatova DE, Baboev SK, Bozorov TA. and Morgunov, A.I. Multivariate Cluster and Principle Component Analyses of Selected Yield Traits in Uzbek Bread Wheat Cultivars. *American Journal of Plant Sciences*,2020; 11, 903-912. https://doi.org/10.4236/ajps.2020.116066

Ahmad, S., Pasha, I., Saeed, M.and Shahid, M.Principal component analysis and correlation studies of spring wheats in relation to cookie making quality*. International Journal of Food Properties*, 2017;20(10), 2299-2313. https://doi.org/10.1080/10942912.2016.1236273

Ambati, D., Phuke, R. M., Vani, V., Sai Prasad, S. V., Singh, J. B., Patidar, C. P. and Dubey, V. G. Assessment of genetic diversity and development of core germplasm in durum wheat using agronomic and grain quality traits. *Cereal Research Communications*, 2020;*48*:375-382.

Atsbeha G, Tesfaye K, Mekonnen T, Haileselassie T. and Kebede M.Genetic diversity and population structure analysis of bread wheat (*Triticum aestivum* L.) germplasms as revealed by inter simple sequence repeat (ISSR) markers. *Genetic Resources and Crop Evolution*.2023; 25:1-5.

Bhandari HR, Bhanu AN, Srivastava K, Singh MN, Shreya HA. Assessment of genetic diversity in crop plants-an overview. Adv. Plants Agric. Res. 2017;7(3):279-86.

Brown-Guedira, G.L., Thompson, J.A., Nelson, R.L. and Warburton, M.L.,. Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Science, 2000;40(3), pp.815-823.

Chaudhary H, Jaiswal JP, Kumar A, Joshi S. Determination of genetic variability and diversity in bread wheat for yield and Yield contributing traits. *International Journal of Plant & Soil Science*. 2022;34(19):16-23.

Elahi, T., Pandey, S. and Shukla, R.S.,. Agro-morphological diversity in promising wheat genotypes grown under restricted irrigated condition. Electronic Journal of Plant Breeding, 2021;12(3), pp.643-651.

FAO. World Food Situation.https://www.fao.org/worldfoodsituation/csdb/e. 2023,Accessed on 15 January 2024

Govindaraj M, Vetriventhan M, Srinivasan M. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives.*Genetics research international*. 2015;2015.

Gower JC. Properties of Euclidean and non-Euclidean distance matrices.Linear algebra and its applications,1985;67:81-97.

Gruet C, Muller D. and Moënne-Loccoz Y. Significance of the Diversification of Wheat Species for the Assembly and Functioning of the Root-Associated Microbiome. *Frontiers in Microbiology*. 2022;12:782135.

Hotelling H. Analysis of a complex of statistical variables into principal components.*Journal of educational psychology*, 1933;24(6):417.

Husson F., Lê S. and Pagès J. Exploratory multivariate analysis by example using R. 2017;CRC press;

Jaiswal, J. P., Arya, M., Kumar, A. and Rawat, R. S. Assessing genetic diversityfor yield and quality traits in indigenous bread wheat germplasm. *Electronic Journal of Plant Breeding*, 2010;1(4): 370-378.

Jaiswal, J.P.. Genetic diversity analysis in exotic germplasm accessions of bread wheat (*Triticum aestivum* L.) by cluster analysis. *Electronic Journal of Plant Breeding*, 2015;6(4):1111-7.

Joshi BK, Ghimire KH, Neupane SP, Gauchan D, Mengistu DK. Approaches and advantages of increased crop genetic diversity in the fields. Diversity. 2023;15(5):603.

Kassambara A. Factoextra: extract and visualize the results of multivariate data analyses. R package version. 2016;1.

Khalid, A., Hameed, A., Shamim, S., & Ahmad, J.divergence in single kernel characteristics and grain nutritional profiles of wheat genetic resource and association among traits. Frontiers in Nutrition, 2022.8.https://doi.org/10.3389/fnut.2021.805446.

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.

Khodadadi M, Fotokian MH, Miransari M. Genetic diversity of wheat (*Triticum aestivum* L.) genotypes based on cluster and principal component analyses for breeding strategies. Australian Journal of Crop Science. 2011;5(1):17-24.

Kumar S, Jacob SR, Mir RR, Vikas VK, Kulwal P, Chandra T, Kaur S, Kumar U, Kumar S, Sharma S, Singh R. Indian wheat genomics initiative for harnessing the potential of wheat germplasm resources for breeding disease-resistant, nutrient-dense, and climate-resilient cultivars. Frontiers in genetics. 2022; 13:834366.

Kumar V, Koutu GK, Singh SK, Mishra DK, Singh PK, Sohgaura N. Genetic analysis of inter subspecific derived mapping population (RILS) for various yield and quality attributing traits in rice. International science journal. 2014;1(3).

Kumar, S., Gupta, V., Yadav, S., Mamrutha, H., Singh, S., Chatrath, R., …and Singh, G. (2021). Multivariate analysis in wheat germplasm captures variability for agro-morphological and physiological traits. The Indian Journal of Agricultural Sciences, 91(9).https://doi.org/10.56093/ijas.v91i9.116079

Le S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. Journal of statistical software. 2008;25:1-8.

Leiva M. Why the price of wheat is so important to food security and global stability. *Investment Mentor.* 2022. Accessed 15 October 2023 .https://www.investmentmonitor.ai/agribusiness/wheat-price-important-food-security-global-stability/?cf-view

Melchinger, A.E., D.R. Buxton, R.A. Forsberg, B.L. Blad, H. Asay, G. M. Paulsen, R.F. Wilson. Use of RFLPmarkers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. Int. crop Sci. 1993;1:621-628.

Mishra, C. N., Tiwari, V., Satish-Kumar, V. G., Kumar, A., & Sharma, I. Genetic diversity and genotype by trait analysis for agro-morphological and physiological traits of wheat (Triticumaestivum L.). *Sabrao J. Breed. Genet*, 2015;*47*(1), 40-48.

Mojena R. Hierarchical grouping methods and stopping rules: an evaluation. The Computer Journal. 1977;20(4):359-63.

Nachimuthu VV, Robin S, Sudhakar D, Raveendran M, Rajeswari S, Manonmani S. Evaluation of rice genetic diversity and variability in a population panel by principal component analysis. Indian journal of science and technology. 2014;10:1555-62.

Pearson K. LIII. On lines and planes of closest fit to systems of points in space. The London, *Edinburgh, and Dublin philosophical magazine and journal of science*. 1901 Nov 1;2(11):559-572.

Peterson, R.F., Campbell, A.B. and Hannah, A.E. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. Canadian journal of research, 1948;26(5), pp.496-500.

Piro, M., Muylle, H., and Haesaert, G. Exploiting rye in wheat quality breeding: the case of arabinoxylan content.*Plants*, 2023; 12(4), 737.https://doi.org/10.3390/plants12040737

Poudel, A., Thapa, D. B., and Sapkota, M. Cluster analysis of wheat (Triticumaestivum L.) genotypes based upon response to terminal heat stress. *International Journal of Applied Sciences and Biotechnology*,2017;*5*(2), 188-193.

Riaz A, Bibi T, Raza Q, Sabar M, Akhter M. Evaluation of rice (*Oryza sativa* L.) advance uniform lines using multivariate analysis. *African Journal of Plant Science*. 2018;12(11):284-189.

Rohlf FJ. NTSYS-pc: numerical taxonomy and multivariate analysis system. Exeter Publishing; 1988.

Saleh, M., Kenaan, R., Suliman, Z., Ali, W., &Jabbour, Y.Diversity analysis and structural modeling for some traits in wheat genotypes. Journal of Phytology, 2021;127-135. https://doi.org/10.25081/jp.2020.v12.6594

Santosh, Jaiswal, J. P., Singh, A. and Gahatyari, N. C. 2019. Genetic diversity analysisin bread wheat (Triticum aestivum L. em. Thell.) for yield and physiological traits.Int. J. Curr. Microbiol. Appl. Sci., 8(2): 3059-3068.

Sarfraz Z, Shah MM, Iqbal MS, Nazir MF, Al-Ashkar I, Rehmani MI, Shahid Iqbal M, Ullah N, El Sabagh A. Rendering Multivariate Statistical Models for Genetic Diversity Assessment in A-Genome Diploid Wheat Population. Agronomy. 2021;11(11):2339.

Sewell M. Principal component analysis. University College London: London, UK. 2008.

Shamuyarira KW, Shimelis H, Figlan S, Chaplot V. Path coefficient and principal component analyses for biomass allocation, drought tolerance and carbon sequestration potential in wheat. Plants. 2022 ;11(11):1407.

Sharma S, Singh V, Mor VS, Punia RC, Hemender H, Khan MU, Sangwan SO. Genetic variability and diversity analysis in wheat (Triticum spp.) genotypes using multivariate techniques.*The Indian Journal of Agricultural Sciences.* 2021;91(11).

Sheela KS, Robin S, Manonmani S. Principal component analysis for grain quality characters in rice germplasm. *Electronic Journal of Plant Breeding.* 2020;11(01):127-31.

Shimelis, Y., Mohammed, W., &Letta, T. (2018).genetic variability and diversity studies in durum wheat (triticumturgidum l. var. durum) genotypes based on cluster analysis using agronomic and quality traits in southern ethiopia. *Journal of plant breeding and Crop Science,*2018;10(5), 99-104.https://doi.org/10.5897/jpbcs2016.0577

Shivramakrishnan, R., Vinoth, R., Arora, A. J. A. Y., Singh, G. P., Kumar, B., & Singh, V. P. Characterization of wheat genotypes for stay green and physiological traits by principal component analysis under drought condition. *Journal of Agricultural Sciences*, 2016;*12*(2), 245-251.

Singh P, Jain PK, Tiwari A. Principal component analysis approach for yield attributing traits in chilli (Capsicum annum L.) genotypes.Chemical Science Review and Letters. 2020;9(33):87-91.

Singh, V. and Singh, A. Analysis of agriculture data using principal component analysis. *International Journal of Multidisciplinary Research and Development*. 2020;7(1):34-37.

Sokal RR, Michener CD. University of Kansas Sci. Bull. 1958;28:1409.

Szczepanik M, Szyszlak-Bargłowicz J, Zając G, Koniuszy A, Hawrot-Paw M, Wolak A. The Use of Multivariate Data Analysis (HCA and PCA) to Characterize Ashes from Biomass Combustion.Energies. 2021;14(21):6887.

Thompson, J.A. and R.L. Nelson. 1998. Utilization of diverse germplasm for soybean yield improvement. *Crop Sci*. 1998;38:1362-1368.

Vus NA, Kobyzeva LN, Bezuglaya ON. Determination of the breeding value of collection chickpea (Cicer arietinum L.) accessions by cluster analysis. Vavilov Journal of Genetics and Breeding. 2020 ;24(3):244.

Yadav SK, Singh AK, Baghel SS, Jarman M, Singh AK. Assessment of genetic variability and diversity for yield and its contributing traits among CIMMYT based wheat germplasm. Journal of Wheat Research. 2014;6(2):154-159.

Zedfar, P. and Golabadi, M. Genetic Variability Assessment in Bread Wheat (Triticum aestivum L.) cultivars under Different Drought Stress Treatments using Multivariate Statistical Analysis. International Journal of Agriculture Innovations and Research. 2013:2(3): 369-372.