PLANT SCIENCE TODAY
Vol x(x): xx–xx
https://doi.org/10.14719/pst.12938

eISSN 2348-1900

RESEARCH ARTICLE

Morphometric diversity and clustering of wild apricot (*Prunus armeniaca* L.) in Himachal Pradesh: Implications for phenotypic selection and germplasm management

Anita Kumari[1], Dinesh Sharma[2], Kehar Singh Thakur[3], Anchal[1*] & P P Sujith[1]

[1]Department of Tree Improvement and Genetic Resources, Dr. YS Parmar University of Horticulture and Forestry, Nauni, Solan 173 230, Himachal Pradesh, India
[2]Directorate of Extension Education, Dr. YS Parmar University of Horticulture and Forestry, Nauni, Solan 173 230, Himachal Pradesh, India
[3]Krishi Vigyan Kendra, Chamba, Dr. YS Parmar University of Horticulture and Forestry, Nauni, Solan 173 230, Himachal Pradesh, India

*Correspondence email - anchkaushal99@gmail.com

Received: 26 November 2025; Accepted: 10 March 2026; Available online: Version 1.0: 20 April 2026

Cite this article: Anita K, Dinesh S, Kehar ST, Anchal, Sujith PP. Morphometric diversity and clustering of wild apricot (*Prunus armeniaca* L.) in Himachal Pradesh: Implications for phenotypic selection and germplasm management. Plant Science Today (Early Access). https://doi.org/10.14719/pst.12938

Abstract

The present study assesses phenotypic variability, trait correlations and clustering patterns among wild apricot populations across 15 sites of Himachal Pradesh during 2013–2014. Thirteen traits related to tree growth, stone and kernel characteristics were recorded from 75 mature trees (5 trees per site). A wide variation was observed in tree growth and kernel traits, with high genetic gain and heritability for the number of branches and diameter at breast height (DBH). Significant correlations existed among important traits, especially between kernel and stone characteristics. Principal component analysis (PCA) revealed that PC1 accounted for 47.78 % variability, mainly influenced by stone and kernel traits. The PC2 and PC3 contributed 21.07 % and 13.04 %, respectively. Cluster analysis grouped sites into 2 main clusters, distinguishing high-performing sites like Ghar and Panjore, valuable for breeding programs focused on specific trait enhancement and oil content. The study effectively identified significant morphometric diversity and strong trait associations in wild apricot populations. Traits like DBH and branch number emerged as prime selection criteria. The PCA and clustering analysis identified superior, trait-rich sites, supporting targeted breeding, improved oil extraction potential, sustainable conservation and utilisation of wild apricot germplasm.

Keywords: cluster analysis; correlation; genetic variability; metric attributes

Introduction

Prunus armeniaca L., commonly known as wild apricot, is a member of the Rosaceae family and one of Southeast Asia's most important stone fruits (1). The wild apricot was originally domesticated in China (2) and it is now cultivated on every continent except Antarctica. *P. armeniaca* is grown in temperate regions worldwide and thrives in certain climates (3). It is a popular fruit of Northern India's hills, which include the majority of Jammu and Kashmir, Himachal Pradesh and the hilly areas of Uttar Pradesh. In Himachal Pradesh, the fruit is typically found in mid-hills, ranging from 1100– 1500 m above mean sea level (4).

The trees of *P. armeniaca* are deciduous, small to medium-sized trees with mature heights of 8–12 m (5). Apricot is a drupe that grows to a width of 3–5 cm, with some types growing up to 8 cm. The fruit's surface is smooth or velvety and yellow or reddish-orange. The seed is centrally placed and protected by a hard shell. The seed and the shell form the stone of the fruit, which has a grainy texture and groovy surface (6). The single seed is enclosed in a hard, stony shell, often called a "stone", with a grainy, smooth texture except for 3 ridges down one side (7). Considerable variation in tree architecture, fruit size, stone and kernel characteristics among wild populations highlights the importance of systematic morphometric evaluation. Quantifying these structural traits provides a basis for assessing phenotypic diversity and identifying distinct population groupings through clustering and multivariate analyses.

Wild apricots are a highly nutritious fruit, containing more carbohydrates, proteins and phosphorus than most other fruits (8). The fruits include antidiarrheal, antipyretic and emetic properties, as well as the ability to quench thirst. However, they are not recommended for the elderly (9). The seeds are used as a tonic and anthelmintic for liver disease, piles, earaches and deafness. In Himachal Pradesh's dry temperate zone of Kinnaur, oil is extracted from the kernels of a wild apricot strain. In these places, wild apricots are also used to manufacture alcoholic beverages (10). As a result of its early maturity, wild apricot cultivation has experienced tremendous growth in recent years.

The development of promising genotypes is one of the most effective ways to enhance the production of wild apricots. The selection is a powerful strategy for developing and aligning a tree adaptation to specific environmental conditions. However, the effectiveness of a selection program is dependent on an accurate estimation of variability, as the genetic and non-genetic components of variation determine the progress of the selection program. Understanding the extent of phenotypic variability among wild

apricot populations is essential for effective characterisation and selection. Partitioning the observed variation into phenotypic and genotypic components provides an initial indication of inheritance patterns, while acknowledging that environmental conditions may influence trait expression under natural field settings.

Furthermore, the genetic variance of any quantitative trait includes both additive and non-additive variance, such as dominance and epistasis. Understanding the heritability of different traits is essential, as only the heritable portion of variation is transmitted from one generation to the next. In addition, a comprehensive understanding of the association of traits is also required for a successful selection program. The association studies between the primary contributing characteristics would be useful in identifying essential marker characters highly associated with expressing specific traits and improving one character without sacrificing the expression of another. Before initiating any tree improvement program, it is necessary to have a thorough understanding of the amount of variability present for various traits among the genotypes. The principal component analysis (PCA) and hierarchical cluster analysis (HCA) are used to categorise individuals and assess variability within and between groups. These methods are useful for visualising and identifying variability through clustering. The PCA is a statistical approach that simplifies complex data by reducing the number of possibly correlated variables into a smaller number of variables called principal components (PC) (11). Hierarchical cluster analysis creates groups based on similarity level using a predetermined metric such as Euclidean distance (12).

Despite the recognised economic and ecological importance of wild apricot in Himachal Pradesh, systematic evaluation of morphometric variability and population structuring remains limited. We hypothesised that substantial phenotypic variability exists among wild apricot populations across different sites, stone and kernel traits contribute significantly to overall variation and multivariate analysis would differentiate phenotypically distinct site-based groups with potential relevance for preliminary selection and germplasm management. Therefore, the present investigation was undertaken to assess phenotypic variability, estimate variability parameters and examine clustering patterns among wild apricot populations based on tree growth, stone and kernel traits.

Materials and Methods

Experimental site and materials

The investigation was carried out during the year 2013–2014 at the Regional Centre NAEB, Nauni, Solan, Himachal Pradesh, India. The survey was conducted in the natural population in 3 regions i.e., Mandi, Sirmour and Solan of Himachal Pradesh. Five sites were selected randomly from each region: Tikkari, Alyas, Bhekhli, Khatchi and Ghar in Mandi; Kharma, Theog, Paoli, Padhar and Matasa in Shimla; and Thyanbag, Chachrarabag, Sangtari, Halan and Panjore in Sirmour (Table 1). Five trees over ten years old were chosen in all *P. armeniaca* plantations at each selected site.

Measurements and observations

The observations on metric traits viz., tree growth, stone and kernel, were recorded. The sampled trees represented phenotypically distinct individuals occurring under natural field conditions; no clonal or pedigree information was available. The tree height and diameter at breast height (DBH) were recorded using the Haga altimeter instrument and tape, respectively. The crown diameter was measured by taking the arithmetic average of the horizontal crown diameter on the north-south axis and on the east-west axis, measured with measuring tape. The number of branches was recorded by visual counting, whereas the length of the clean bole was recorded using a measuring stick. Stones from the identified trees were collected and kernels were separated from the mass of the stone by using the specific gravity separation method after breaking the stones with a mechanical decorticator. The stone and kernel traits were measured in millimeters using a digital caliper.

The coefficient of variability was calculated using the methods recommended by previous researchers (13, 14) i.e.,

$$\text{Genotypic coefficient of variability [GCV (\%)]} = \frac{\sqrt{V_g}}{\bar{X}} X\ 100$$

$$\text{Phenotypic coefficient of variability [PCV (\%)]} = \frac{\sqrt{V_p}}{\bar{X}} X\ 100$$

Where,

Vg = genotypic variance

Vp = phenotypic variance

$\bar{X}$ = trait mean

Since the study was conducted under natural site conditions without controlled progeny or clonal replication, the estimated variance components primarily represent phenotypic variation and should be interpreted cautiously as indicative rather than definitive measures of genetic parameters.

$$\text{gain [GG (\%)]} = \frac{GA}{\bar{X}} X\ 100$$

Heritability in the broad sense expected genetic advance at 5 % selection intensity and genetic gain was calculated as indicated by earlier researchers (13, 15, 16). Broad-sense Genetic

Table 1. Altitude, latitude and longitude of the sites selected for study

Region	Sites	Altitude (m)	Latitude (°N)	Longitude (°E)
	Tikkari	850 to 980	31.41 to 31.66	76.80 to 76.95
	Alyas	760 to 864	31.42 to 31.70	76.55 to 76.93
Mandi	Bhekhli	1035 to 1210	31.15 to 31.26	76.39 to 76.48
	Khatchi	1300 to 1520	31.77 to 31.92	76.89 to 77.02
	Ghar	1920 to 2010	31.34 to 31.56	77.33 to 77.43
	Kharma	1550 to 1620	31.23 to 31.33	77.44 to 77.61
	Theog	1985 to 2123	31.12 to 31.34	77.35 to 77.40
Shimla	Paoli	1560 to 1730	31.06 to 31.33	77.13 to 77.18
	Padhar	1222 to 1445	30.23 to 30.56	75.48 to 75.86
	Matasa	1245 to 1366	31.02 to 31.55	77.23 to 77.35
	Thyanbag	560 to 670	30.22 to 30.56	77.05 to 77.51
	Chachrarabag	1364 to 1512	30.22 to 30.68	77.01 to 77.23
Sirmour	Sangtari	998 to 1107	30.34 to 30.42	77.08 to 77.12
	Halan	1524 to 1655	30.72 to 30.88	77.62 to 77.67
	Panjore	1750 to 1956	30.24 to 30.87	77.09 to 77.23

$$\text{heritability } [H^2_{BS}(\%)] = \frac{V_g}{V_p} X\ 100$$

$$\text{Genetic advance } [GA] = H^2_{BS} X \sigma_p X K$$

Where;

σ_p = standard deviation of phenotypic variance

K = selection differential (2.06 at 5 % selection intensity)

Statistical analysis

The statistical mean was determined using the previous approach (17). The "R"-based software package "metan" was used to estimate the correlation coefficient (18). The “factoextra” (19) and "FactoMineR" packages were used for principal component and cluster analysis (20).

Since observations were recorded from naturally grown trees across different sites, the estimated variance components primarily represent phenotypic variation influenced by both genetic and environmental factors.

Results and Discussion

Estimates of range in variation

The variation for tree height, DBH, crown diameter, number of branches and length of clean bole ranged from 9.86–12.66 m, 81.20–296.00 cm, 5.80–10.29 m, 3.80–22.80 and 0.80–2.60 m, respectively. The stone length, width, thickness and weight of 100 stones varied from 18.42–23.29 mm, 14.60–18.94 mm, 9.90–12.76 mm and 1068.76 –2067.60 g, respectively. The wide range in kernel length (13.42–16.60 mm), kernel width (8.21–11.10 mm), kernel thickness (4.69–6.43 mm) and weight of 100 kernels (324.40–538.40 g) indicated the ample scope for trait-specific extraction of oil in the future (Table 2).

Phenotypic and genotypic coefficient of variation

Variability is an important consideration when selecting desirable genotypes. The phenotypic coefficients of variation (PCV) and genotypic coefficients of variation (GCV) were used to analyse genotype variability. Respective values of the highest phenotypic and genotypic coefficient of variation were estimated for the number of branches (58.12 %, 48.32 %), DBH (46.44 %, 36.76 %) and length of clean bole (45.58 %, 21.80 %); moderate for crown diameter (24.97 %, 16.17 %) and tree height (22.06 %, 10.19 %). While other traits such as weight of 100 stones (15.92 %, 8.83 %), stone width (15.24 %, 11.70 %), kernel thickness (15.16 %, 12.47 %), weight of 100 kernels (14.63 %, 5.34 %), kernel width (12.96 %, 10.02 %), stone thickness (12.80 %, 9.76 %), stone length (12.33 %, 9.41 %) and kernel length (12.23 %, 9.66 %) exhibited low level of phenotypic and genotypic coefficient of variation, respectively. All characters had a slightly greater phenotypic coefficient of variation than their corresponding genotypic coefficients, suggesting that apparent variation is influenced by both genotype and environment. However, the proximity between PCV and GCV should not be interpreted as definitive evidence of environmental stability, as variance components were derived from site-based observations without controlled replication. Therefore, these estimates reflect relative variability patterns under natural field conditions rather than precise partitioning of genetic and environmental effects (21).

Heritability estimates in a broad sense and genetic gain

The heritability of several traits ranged between 13.32 to 69.14 %, indicating low to high inheritance. However, the genetic gains were noted as low to high (4.01 to 82.78 %) (Table 2). The high estimates of heritability with high genetic gain were noted for the number of branches and DBH. The high heritability of these traits may be attributed to their polygenic but structurally stable nature, as vegetative growth parameters are often under strong genetic regulation and less prone to short-term environmental fluctuations compared to reproductive traits. Similar trends have been reported in apricot and other temperate fruit species (22, 23), where growth-related traits showed relatively higher heritability, suggesting a greater potential contribution of additive gene effects under phenotypic evaluation. The estimates of heritability were high, with a moderate percentage of genetic gain for kernel thickness. However, the high heritability with a low genetic gain was realised for kernel length, kernel width, stone length, stone width and stone thickness. This pattern is often suggested as limited additive variance, such inference should be made cautiously in the absence

Table 2. Estimates of variability for different metric attributes of wild apricot

Traits	Range	Mean±SE(m)	Coefficients of variability (%)		Heritability (%)	Genetic advance	Genetic gain (%)
			Genotypic	Phenotypic			
Tree height (m)	9.86-12.66	10.48±0.96	10.19	22.06	21.33	1.02	9.69
Diameter at breast height (cm)	81.20-296.00	147.32±39.14	36.76	46.44	62.66	88.31	59.94
Crown diameter (m)	5.80-10.29	7.95±0.99	16.17	24.97	41.93	1.71	21.57
Number of branches	3.80-22.80	12.41±4.28	48.32	58.12	69.14	10.28	82.78
Length of clean bole (m)	0.80-2.60	1.62±0.31	21.80	45.58	22.88	0.35	21.48
Stone length (mm)	18.42-23.29	20.71±1.02	9.41	12.23	59.27	3.09	14.93
Stone width (mm)	14.60-18.94	16.47±1.02	11.70	15.24	58.92	3.05	18.50
Stone thickness (mm)	9.90-12.76	10.84±0.57	9.76	12.80	58.09	1.66	15.32
Weight of 100 stones (g)	1068.76-2067.60	143.17±11.99	8.83	15.92	30.79	14.45	10.09
Kernel length (mm)	13.42-16.60	14.49±0.69	9.66	12.23	62.39	2.28	15.72
Kernel width (mm)	8.21-11.10	9.61±0.50	10.02	12.96	59.83	1.53	15.97
Kernel thickness (mm)	4.69-6.43	5.69±0.31	12.47	15.16	67.67	1.20	21.13
Weight of 100 kernels (g)	324.40-538.40	39.77±3.43	5.34	14.63	13.32	1.60	4.01

of controlled experimental design capable of reliably separating additive, dominance and environmental components.

Heritability in the broad sense includes both additive and non-additive gene action, but it is the additive portion of the total variance that is responsive to selection. Therefore, broad-sense heritability estimates alone may not be a useful index for predicting a character in an improvement program. The selection appears to be more effective for a character with high heritability and high genetic gain, indicating lesser environmental influence and greater contribution of additive gene action in determining inheritance. A moderate-to-low heritability accompanied by a moderate-to-low genetic gain indicates non-additive genes are at play in the expression of these traits and recombination breeding may provide a better tool for exploiting them (24). The proportion of genetic and non-genetic components of variation determines the progress in terms of genetic gain in breeding programs. The study of heritability and genetic gain is highly valuable in determining the scope for improvement through selection since many of the economic characteristics, including the weight of 100 stone and kernels, are complex and are heavily influenced by environmental variables. Previous researchers studied the genetic variability and association of component characters for fruit weight in apricot cultivars (25).

It is important to note that broad-sense heritability estimated from site-based morphometric data may be confounded by environmental heterogeneity. Therefore, the present estimates should be interpreted as indicators of phenotypic inheritance potential rather than definitive measures of genetic transmissibility.

Character association

Character association measures the relationship between pairs of traits in terms of the correlation coefficient. This association allows the identification of traits that may serve as useful indicators for phenotypic selection and preliminary improvement strategies (26). The present study reveals that the maximum, positive and highly significant correlation was found between the weight of 100 stones and the weight of 100 kernels i.e., 0.94, whereas the number of branches and crown diameter were highly negatively and non-significantly correlated (Fig. 1).

The tree height (0.52) and crown diameter (0.77) showed a significant positive correlation with DBH. The number of branches exhibits a significant positive correlation with the weight of 100 stones (0.62), the weight of 100 kernels (0.62), stone thickness (0.57) and kernel length (0.55). The length of the clean bole showed a significant positive correlation with kernel thickness (0.61). The stone length had a significant positive correlation with stone width (0.93), weight of 100 stones (0.87), kernel length (0.85), kernel width (0.80), weight of 100 kernels (0.70) and stone thickness (0.69). The stone width is highly correlated with the weight of 100 stones (0.93), the weight of 100 kernels (0.86), kernel length (0.81), kernel width (0.77) and stone thickness (0.77). The stone thickness showed a significant positive correlation with the weight of 100 stones (0.84), the weight of 100 kernels (0.84) and kernel length (0.62). The weight of 100 stones showed a significant positive correlation with kernel length (0.84) and kernel width (0.74). The kernel length had a significant positive correlation with the weight of 100 stones (0.84) and the weight of 100 kernels (0.81).

These results indicate coordinated phenotypic variation among fruit, stone and kernel traits. However, correlation reflects statistical association and does not necessarily imply direct causal relationships or predictive breeding value, especially under survey-based filed conditions where environmental factors may influence multiple traits simultaneously. Therefore, the identified associations should be interpreted as indicative patterns of co-variation that may guide further investigation. Validation through controlled experiments or progeny evaluation would be necessary before translating these associations into definitive breeding strategies. Similar positive associations have been reported by previous researchers as they reported a positive and significant correlation between kernel weight and different stone and kernel characters of wild apricot (27). Correlation studies in wild apricot plus trees were reported earlier (28). They found that the kernel thickness and kernel weight were positively and highly significantly correlated with the stone thickness.

Principal component analysis

The PCA reduced the original data set of thirteen metric attributes to eleven independent vector or principal components (PCs) that had a cumulative explained variance of 100 %. The Eigenvalues and associated percentage of variation of the respective principal components (PCs) are presented in Table 3.

The first 3 components in the PCA analysis with Eigenvalues greater than one contributed 81.90 % of the total variability. The PC1 with an Eigenvalue of 6.21 accounts for the maximum variability in the data set (47.78 %), while PC2 with an Eigenvalue of 2.73 accounts for 21.07 %. The PC3 had an Eigenvalue of 1.69 and contributed 13.04 % to the observed variability. Only the first 3 principal components (PCs) with the Eigenvalues >1 explaining 81.90 % of the variations among the population were chosen for further analysis.

Table 3. Principal component analysis of the studied wild apricot trees

Characters	PCI	PCII	PCIII
Tree height (m)	-0.11	0.37	0.17
Diameter at breast height (cm)	0.06	0.48	0.16
Crown diameter (m)	0.09	0.52	0.18
Number of branches	0.22	-0.32	-0.19
Length of clean bole (m)	0.03	0.35	-0.50
Stone Length (mm)	0.36	0.10	0.17
Stone Width (mm)	0.38	0.05	0.14
Stone Thickness (mm)	0.34	-0.02	-0.28
Weight of 100 stones (g)	0.39	-0.06	0.02
Kernel Length (mm)	0.35	-0.15	0.20
Kernel Width (mm)	0.32	0.17	0.00
Kernel Thickness (mm)	0.04	0.19	-0.65
Weight of 100 kernels (g)	0.37	-0.05	-0.07
Eigenvalues	6.21	2.73	1.69
Per cent of variability	47.78	21.07	13.04
Cumulative per cent of variability	47.78	68.85	81.90

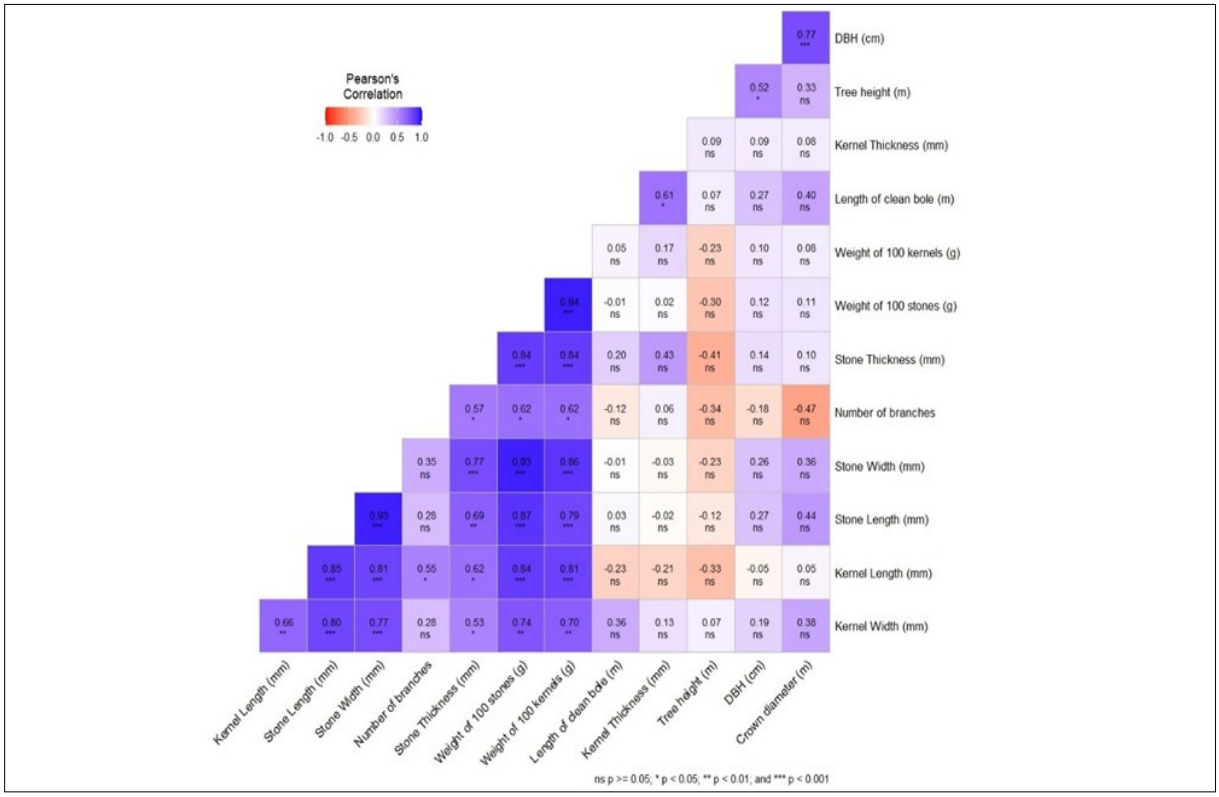

Fig. 1. Correlation values for morphological, stone and kernel traits of wild apricot.

It is important to note that several stone and kernel traits are biologically interrelated (e.g., stone size, kernel size and their respective weights). Such structural interdependence can lead to high correlation coefficients and may disproportionately influence the variance captured by PC1. Therefore, the strong contribution of stone and kernel traits to the first principal component likely reflects coordinated phenotypic scaling rather than entirely independent sources of variation. This inherent trait interdependence should be considered when interpreting PCA results, as multicollinearity among related traits can amplify their contribution to overall variability patterns.

The major contributing character in the principal component one (PC1) include weight of 100 stones (0.39) followed by the stone width (0.38), weight of 100 kernels (0.37) and stone length (0.36) while, tree height (-0.11) had the highest negative loading. In the PC2, the highest positive loading was obtained from crown diameter (0.52) followed by DBH (0.48) and tree height (0.37); and the maximum negative loading was due to the number of branches (-0.32) and kernel length (-0.15). Moreover, in PC3, a trait such as kernel length (0.20) had the highest positive loading; and kernel thickness (-0.65), length of clean bole (-0.50) and stone thickness (-0.28) exhibited maximum negative loading. Positive and negative loading of factors were observed in the above-discussed major principal components, which indicates that the components and variables had positive and negative correlations.

The biplot analysis (Fig. 2) provides a graphic representation of the relationship among the traits and the magnitude to which the traits contributed to diversity. The important consideration of the biplot diagram is the angles of vectors. An acute angle demonstrates a positive correlation, while an obtuse angle displays a negative correlation and a right angle displays no correlation at all. A high positive correlation has been observed between the weight of 100 kernels and weight of 100 stones, stone thickness, stone length, stone width, kernel width and kernel length. Biplot analysis also revealed sites namely Ghar, Sangtari, Chachrarabag, Thyanbag and Panjore to be completely distinguishable from the other sites which performed well for the PC1 and PC2 (Fig. 3). The separation is based on the higher levels of DBH, weight of 100 stones, stone width, tree height, weight of kernels, stone length, kernel length and stone thickness of respective sites.

Principal component analysis is widely used in tree improvement to reduce the number of variables and to group genotypes based on their similarities or differences. The biplot highlights the traits that contribute most to the overall diversity among wild apricot genotypes. Traits like weight of 100 kernels, weight of 100 stones, stone thickness, stone length, stone width, kernel width and kernel length are strongly correlated and play a significant role in differentiating genotypes. The present study conforms with those of (27, 29, 30–32).

The PCA and biplot analyses were used as exploratory tools to summarise trait relationships and phenotypic diversity patterns among sites. The identified trait grouping and site differentiation should therefore be interpreted descriptively rather than as confirmatory evidence of genetic divergence.

Hierarchical cluster analysis

Prior to multivariate analysis, all morphometric variables measured in different units (m, cm, mm, g) were standardised using z-score transformation (mean = 0, standard deviation = 1) to eliminate scale effects and ensure equal contribution of traits to distance calculations. HCA was then performed using Euclidean distance to assess phenotypic dissimilarity among sites. All the studied sites were grouped into 2 main clusters as per the data depicted in Fig. 4. Cluster I consists of five sites, i.e., Panjore, Chachrarabag, Thyanbag and Sangtari, of the Sirmour region and the Ghar sites of the Mandi region. Cluster II was further bifurcated into two sub-clusters, i.e., sub -cluster IIa and sub-cluster IIb. The sub-cluster IIa consists of the Kharma, Theog and Matasa sites of the Shimla region, whereas the rest were placed into sub-cluster IIb. The sites belonging to the same

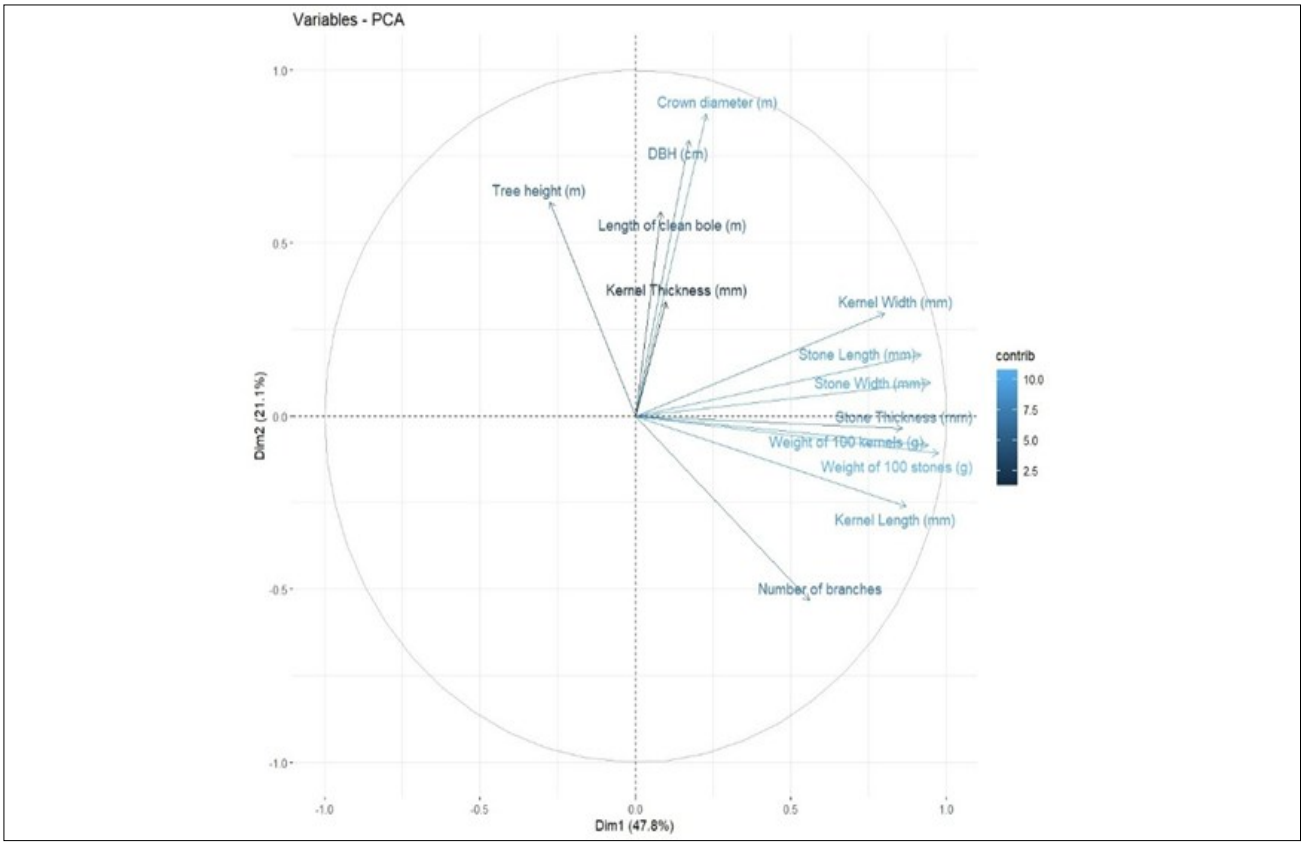

Fig. 2. Character contribution for genetic divergence for the PC1 and PC2.

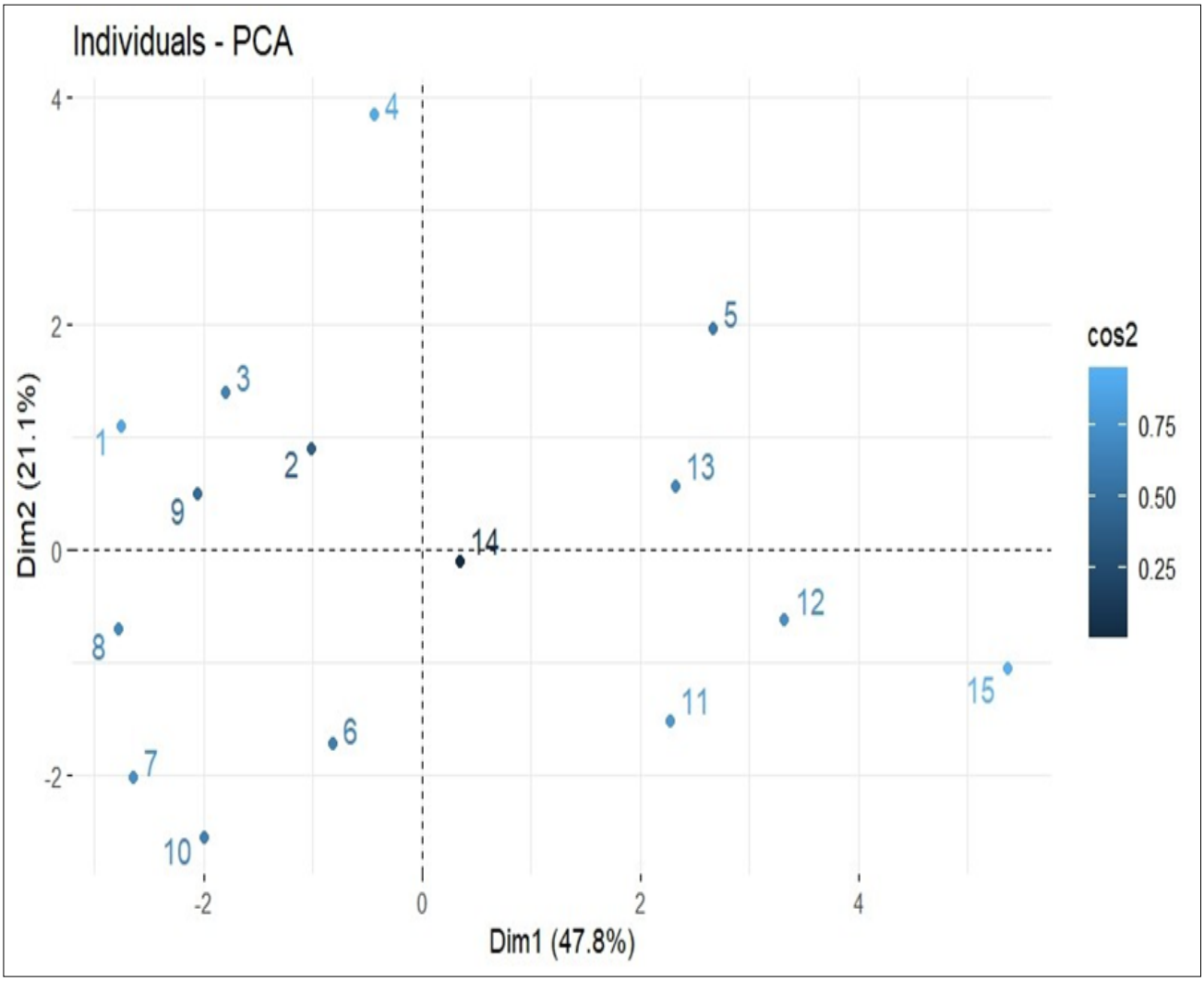

Fig. 3. Distribution of genotypes in the scatter plot along with PC1 and PC2. Where, 1: Tikkari; 2: Alyas; 3: Bhekhli; 4: Khatchi; 5: Ghar; 6: Kharma; 7: Theog; 8: Paoli; 9: Padhar; 10: Matasa; 11: Thyanbag; 12: Chachrarabag; 1 3: Sangtari; 14: Halan; 15: Panjore.

regions were classified under the same cluster, indicating limited gene flow across regions and region-specific trait combinations.

From a breeding perspective, cluster analysis provides a strategic framework for parent selection. Genotypes belonging to different clusters are expected to be more phenotypically divergent and crossing such parents can enhance heterosis and generate broader variability in segregating populations. For instance, superior sites like Ghar and Panjore from Cluster I, which exhibited desirable growth and kernel traits, can be crossed with genetically distant genotypes from sub-cluster IIb to continue to combine complementary traits such as higher oil content, kernel size or tree vigour. Inter-cluster hybridisation is particularly valuable for accumulating favourable additive genes while also exploiting non-additive gene effects. In contrast, intra-cluster selection may be useful for trait stabilisation and pure-line selection where uniformity is desired. Thus, the identified clusters not only reveal genetic relationships but also serve as a practical guide for designing cross-breeding programs aimed at trait enhancement, oil yield improvement and long-term genetic conservation of wild apricot germplasm.

The clustering of wild apricot sites into two main groups and subsequent sub-clusters reveals distinct phenotypic groupings. This information is crucial for understanding the evolutionary history and genetic diversity of the wild apricot. Our results reflect the findings of previous researchers for cluster analysis (29). They generated a dendrogram based on UPGMA distance and classified 20 genotypes of apricots into 4 main clusters. Like our findings, 92 wild apricot genotypes were classified into two main clusters based on 24 traits (32).

These clustering patterns provide a useful phenotypic framework for parent selection, which should be further validated through molecular marker analysis or progeny testing before being applied in advance breeding programs.

Conclusion

The current study revealed substantial phenotypic variability and important trait associations among wild apricot populations across Himachal Pradesh. Growth traits such as diameter at breast height and number of branches exhibited relatively high heritability and genetic gain, indicating their potential usefulness as selection criteria in improvement programs. Multivariate analysis, including PCA and hierarchical clustering, effectively differentiated trait-rich sites and highlighted promising populations for targeted breeding and oil-related utilisation. An understanding of heritability remains essential, as only the inheritable portion of trait variation is transmitted across generations and contributes to selection response. However, since the present investigation was based primarily on morphometric evaluation, the observed variability reflects phenotypic investigation was based primarily on morphometric evaluation, the observed variability reflects phenotypic expression influenced by both genetic and environmental factors. Therefore, the findings provide a valuable preliminary framework for selection and germplasm utilisation, which should be further validated through molecular characterisation or progeny-based studies for more precise estimation of genetic divergence. The study contributes useful insights into phenotypic diversity patterns and offers practical guidance for breeding, conservation and sustainable utilization of wild apricot resources.

Acknowledgements

The authors are highly thankful to the NAEB (National Afforestation and Eco Development Board) Research Centre, Dr. YS Parmar University of Horticulture and Forestry, Nauni, Solan, Himachal Pradesh, India, for providing facilities and support during the work.

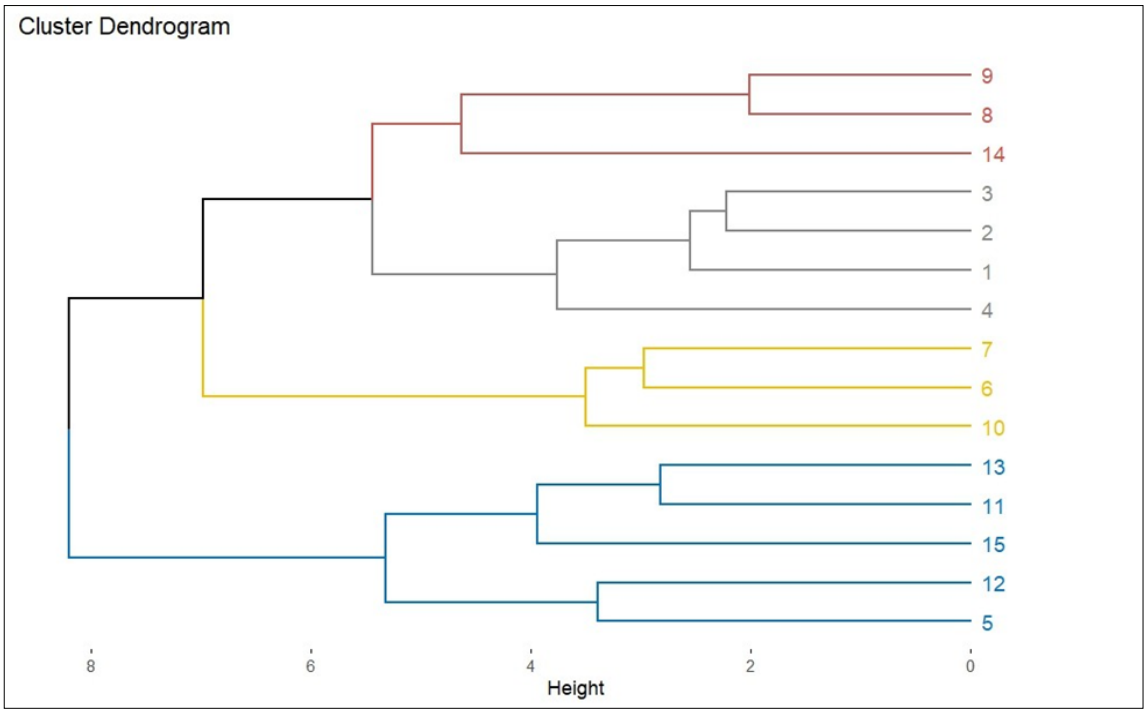

Fig. 4. Dendrogram clustering of fifteen sites of wild apricot. Where, 1: Tikkari; 2: Alyas; 3: Bhekhli; 4: Khatchi; 5: Ghar; 6: Kharma; 7: Theog; 8: Paoli; 9: Padhar; 10: Matasa; 11: Thyanbag; 12: Chachrarabag; 13: Sangtari; 14: Halan; 15: Panjore.

Authors' contributions

A carried out data collection, conceptualisation, review and editing. DS carried out the investigation, review and editing and provided supervision. PPS participated in formal analysis. AK drafted the manuscript and performed data analysis. AK and KST cross-verified the data. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interest to declare.

Ethical issues: None

References

1. Ali S, Masud T, Abbasi KS. Physico-chemical characteristics of apricot (*Prunus armeniaca* L.) grown in northern areas of Pakistan. Sci Hortic. 2011;130(2):386–92. https://doi.org/10.1016/j.scienta.2011.05.040
2. Liu S, Cornille A, Decroocq S, Tricon D, Chague A, Eyquard JP, et al. The complex evolutionary history of apricots: species divergence, gene flow and multiple domestication events. Mol Ecol. 2019;28(24):5299–314. https://doi.org/10.1111/mec.15296
3. Fadón E, Herrera S, Guerrero BI, Guerra ME, Rodrigo J. Chilling and heat requirements of temperate stone fruit trees (*Prunus* sp.). Agronomy. 2020;10(3):409. https://doi.org/10.3390/agronomy10030409
4. Sharma R, Gupta A, Abrol GS, Joshi VK. Value addition of wild apricot fruits grown in north-west Himalayan regions: a review. J Food Sci Technol. 2014;51:2917–24. https://doi.org/10.1007/s13197-012-0766-0
5. Li W, Liu L, Wang Y, Zhang Q, Fan G, Zhang S, et al. Genetic diversity, population structure and relationships of apricot (*Prunus*) based on restriction site-associated DNA sequencing. Hortic Res. 2020;7:69. https://doi.org/10.1038/s41438-020-0284-6
6. Leccese A, Bartolini S, Viti R. Total antioxidant capacity and phenolics content in apricot fruits. Int J Fruit Sci. 2007;7(2):3–16. https://doi.org/10.1300/J492v07n02_02
7. Lingdi L, Bartholomew B. Armeniaca. In: Wu ZY, Raven PH, editors. Flora of China. Vol. 9. Beijing: Science Press; 2003. p.396–401.
8. Fratianni F, Ombra MN, d'Acierno A, Cipriano L, Nazzaro F. Apricots: biochemistry and functional properties. Curr Opin Food Sci. 2018;19:23–9. https://doi.org/10.1016/j.cofs.2017.12.006
9. Pawar KR, Nema PK. Apricot kernel characterization, oil extraction and its utilization: a review. Food Sci Biotechnol. 2023;32(3):249–63. https://doi.org/10.1007/s10068-022-01228-3
10. Pal DK, Dutt B, Attri RD. Studies on important medicinal plants of mid Himalayan region of Himachal Pradesh. Int J Pharm Pharm Sci. 2020;9:158–75.
11. Ziegel E. Encyclopedia of environmetrics. Technometrics. 2002;44:408–9. https://doi.org/10.1198/tech.2002.s97
12. Richard AJ, Dean WW. Applied multivariate statistical analysis. London: Prentice Hall; 2002.
13. Burton GW, DeVane EW. Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. Agron J. 1953;45:478–81. https://doi.org/10.2134/agronj1953.00021962004500100005x
14. Pillai SK, Sinha HC. Statistical methods for biological workers. Agra: Ram Prasad and Sons; 1968. p.610.
15. Lush JC. Intersire correlations and regression of offspring on dams as a method of estimating heritability. Proc Am Soc Anim Prod. 1940;33:293–301. https://doi.org/10.2527/jas1940.19401293x
16. Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in soybeans. Agron J. 1955;47:314–8. https://doi.org/10.2134/agronj1955.00021962004700070009x
17. Goulden CH. Methods of statistical analysis. 2nd ed. New York: Wiley; 1952.
18. Olivoto T, Lúcio ADC. metan: an R package for multi-environment trial analysis. Methods Ecol Evol. 2020;11(6):783–9. https://doi.org/10.1111/2041-210X.13384
19. Kassambara A. factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0; 2016. https://doi.org/10.32614/CRAN.package.factoextra
20. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008;25:1–18. https://doi.org/10.18637/jss.v025.i01
21. Singh R, Jalikop SH. Studies on variability in grape. Indian J Hortic. 1986;47:207–9.
22. Srivastava KK, Sharma AK. Genetic variability, heritability and genetic advance relating to physico-chemical properties of apricot (*Prunus armeniaca* L.) genotypes in Kashmir valley. Haryana J Hortic Sci. 2006;35(1–2):41.
23. Karaat FE, Serce S. Heritability estimates and variation of pomological traits, total phenolic compounds and antioxidant capacity in two apricot progenies. Turk J Agric For. 2020;44(1):54–61. https://doi.org/10.3906/tar-1811-66
24. Dudhe MY, Mulpuri S, Meena HP, Ajjanavara RRG, Kodeboyina VS, Adala VR. Genetic variability, diversity and identification of trait specific accessions from conserved sunflower germplasm for exploitation in breeding programme. Agric Res. 2019;9:1–14. https://doi.org/10.1007/s40003-019-00406-w
25. Sofi AA, Zaffar G, Mir MS. Genetic variability and association of component characters for fruit weight in apricot (*Prunus armeniaca* L.) cultivars of Kargil (Ladakh). Indian J Hortic. 2001;58(3):239–43.
26. Mesera E, Shifaraw G, Alamerew S, Amsalu B. Genetic variability analysis and association of traits in common bean (*Phaseolus vulgaris* L.) landraces collected from Ethiopia at Jimma. Adv Agric. 2022;2022:1–14. https://doi.org/10.1155/2022/4400711
27. Singh NB, Chaudhary VK. Variability, correlation and path analysis between kernel yield and other nut characters in wild apricot. Adv Hortic For. 1993;3:60–7.
28. Kumar A, Bhan S. Correlation studies in wild apricot (*Prunus armeniaca* L.) plus trees. J Hortic For. 2010;2(2):17–21.
29. Mratinić E, Popovski B, Milošević T, Popovska M. Analysis of morphological and pomological characteristics of apricot germplasm in FYR Macedonia. J Agric Sci Technol. 2011;13:1121–34.
30. Milosevic T, Milosevic N, Glisic I, Glisic IS. Determination of size and shape properties of apricots using multivariate analysis. Acta Sci Pol Hortorum Cultus. 2014;13:77–90.
31. Ayour J, Sagar M, Harrak H, Alahyane A, Alfeddy MN, Taourirte M, et al. Evolution of some fruit quality criteria during ripening of twelve new Moroccan apricot clones (*Prunus armeniaca* L.). Sci Hortic. 2017;215:72–9. https://doi.org/10.1016/j.scienta.2016.12.010
32. Chauhan N, Singh D, Kumar K, Dogra RK. Genetic variability, character association and diversity studies on wild apricot (*Prunus armeniaca* L.) genotypes in Himachal Pradesh, India. Genet Resour Crop Evol. 2020;67:1695–705. https://doi.org/10.1007/s10722-020-01003-5

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics,

NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.