Journal of Animal & Plant Sciences, 33(6): 2023, Page: 1346-1355
ISSN (print): 1018-7081; ISSN (online): 2309-8694
https://doi.org/10.36899/JAPS.2023.6.0674

GENETIC DIVERSITY OF THE BC_2 POPULATION OF ALFALFA MULTIFOLIATE LEAVES BASED ON MORPHOLOGICAL TRAITS USING CORRELATION, PRINCIPAL COMPONENT, AND CLUSTERING ANALYSIS

O. I. M. Olom[1, 2], Z. Wei[1*] and L. Nana[1]

[1]Institute of Grassland Science, College of Animal Science and Technology, Yangzhou University, Yangzhou 225009, China.
[2]Department of Forestry and Range Science, Faculty of Natural Resources and Environmental Studies, University of Kordofan, Elobeid, Sudan.
[*] Corresponding author's Email: zwwei@yzu.edu.cn

ABSTRACT

The current research aimed to assess the 179 BC_2 population of alfalfa (*Medicago sativa* L. cv. Huaiyin) and PL34HQ multifoliate leaves and their parents to select genotypes based on morphological traits for future alfalfa breeding programs. A backcross breeding was carried out between the recurrent male parent and alfalfa PL34HQ as the non-recurrent female parent of alfalfa. In the pod setting stage, morphological traits were measured in the BC_2 populations. Data were analyzed for descriptive statistics, correlation, principal component (PCA), and agglomerative hierarchical cluster analyses. The result of descriptive statistics revealed high variation in the means, minimum, and maximum values for most of the morphological traits. Therefore, high variability was observed in plant height (PH) 26-98 cm, number of branches per plant (NBP) 6-74, stem diameter (SD) 1.46-5.46 mm, second branch length SBL 0-62 cm, number of multifoliate leaves per plant (NMFLP) 0-95.7%, number of multifoliate leaves on the main branch (NMFLMB) 0-90.6%, number of nodes (NN) 12-24, fresh shoot weight (FSW) 3.10-359.4 g, and dry shoot weight (DSW) 0.40-108.6 g. The results showed that the comprehensive traits of the BC_2 population tended to the recurrent parent, about 32.40% of the total population. However, many plants tended to the non-recurrent parent, about 67.40% of the total population, with a high multifoliate rate. Correlation analysis revealed a significant positive association of plant height with stem diameter, fresh shoot weight, dry shoot weight, second branch length, number of nodes, and branch number. Principal component analysis (PCA) discovered the first three components with an eigenvalue greater than one, contributing 78.98% of total variability for different traits. The PC1 showed favorable factor loading for all morphological traits. The contribution of dry shoot weight, fresh shoot weight, stem diameter, and plant height was highest in principal component 1. The average hierarchical cluster analysis method grouped all the genotypes into four main clusters. Alfalfa multifoliate leaf genotype clustering based on morphological traits would help identify potential genotypes for future breeding research and programs.

Keywords: - Alfalfa multifoliate leaves, Principal component analysis, Correlation, Cluster analysis.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

Published first online August20, 2023 Published final December 13, 2023

INTRODUCTION

Alfalfa (*Medicago sativa* L.) is a globally important leguminous fodder crop because of its high production and nutritional value (Adhikari *et al*., 2018). Alfalfa has a high protein content, making it excellent hay and pasture crop for animals, particularly dairy cows (Riday and Brummer, 2002; Li and Brummer, 2009). Alfalfa is autotetraploid (2n = 4× = 32), long selection cycles, tetrasomic inheritance, significant inbreeding depression, and genotype-environment interaction have hampered alfalfa yield improvement (Brummer, 1999; Annicchiarico *et al*., 2010). Recurrent behavioral selection is used in most alfalfa breeding projects to enhance yield and stress tolerance. However, new breeding methods and innovative breeding approaches have been used lately (Li and Brummer, 2012; Biazzi *et al*., 2017).

Alfalfa is a perennial plant that is mainly cross-pollinated; several variables influence self-fertility. The natural rate of cross-pollination for alfalfa ranges from 25% to 75%. Physiological and morphological traits limit alfalfa self-pollination, producing low selfing rates. As an autotetraploid, it requires different breeding methods and offspring selection than diploid plant species. Because of these factors, the use of some new breeding technology in alfalfa for major food crops has been delayed.

Selective breeding is the process of selecting cultivars and genotypes with desirable traits to develop new cultivars. It's important to improve existing varieties and breed new cultivars from germplasm collections or other sources. In China, alfalfa is grown through selective breeding, backcrossing, breeding of male-sterile lines, breeding in space, and breeding with the help of biotechnology (Shi *et al*., 2017).

Cross-pollination is used for breeding varieties by choosing hybrid progeny. It may recombine the parents' genes to produce different varieties and provide selection material. Genetic recombination lets parents with different major characteristics contribute desirable alleles. Selecting the right parents and making the right combinations during crossbreeding is critical. Currently, the majority of alfalfa varieties recognized in China were developed through the use of traditional breeding techniques. These simple approaches take advantage of excellent germplasm resources in various ecological regions to develop new varieties adapted for regional conditions. However, these procedures take a long time to develop farm varieties. The application of genetic engineering and other modern biotechnology in alfalfa breeding has been put in place in China for only 20 years.

Multifoliate (MF) alfalfa cultivars characterized by four or more leaflets per leaf than three have been marketed for greater nutritional value and intake potential than standard trifoliate (TF) alfalfa cultivars. Increasing the ratio of leaves to stems is one of the ways to improve the quality of forage in alfalfa, and multifoliate genotypes are considered potential in this respect. Breeding for high expression of multifoliate character effectively reduces neutral and acid detergent fiber concentrations in green mass by increasing the number of leaves without decreasing dry matter yield (Juan *et al*., 1993). Leaves are essential for photosynthesis, thermoregulation, and gas exchange. Plant leaf shapes are amazingly different and can be classified into simple and compound leaves. A morphogenetic process gives compound leaves greater variation than simple leaves. Therefore, one of the goals of alfalfa breeding is the establishment of multifoliate cultivars, which enhance the quality of alfalfa fodder mass, foliage, and other essential forage indicators (Bingham and Murphy, 1965).

Multivariate statistical methods are required for classifying germplasm and investigating genetic correlations among breeding programs. Among multivariate approaches, cluster and principal component analysis are desirable tools for the morphological characterization of genotypes and their clustering according to similarity (Khalid *et al*., 2022). Cluster and principal component analysis (PCA) help identify genotypes from distant clusters and select essential characteristics contributing to the total variance. The principal component analysis was used to compare classification factors before cluster analysis. Cluster analysis is a useful technique for determining familial relationships, i.e., calculating the genetic affinity or distance between genotypes (Mishra *et al*., 2015). In addition, cluster Analysis helps breeders select parents by grouping genotypes based on weighted means (Dangi *et al*., 2018).

The study of genetic differences can help choose which genotypes to use in breeding programs to make new populations because it measures the level of diversity among the selected genotypes (Khan *et al*., 2020). Many genetic divergence studies have been done on different crop species based on qualitative and quantitative traits to find genetically different parents to use in hybridization (Khalid and Hammed, 2021).

This investigation was aimed to evaluate genetic diversity among alfalfa multifoliate leaves to select genotypes based on the best morphological traits for further alfalfa breeding programs.

MATERIALS AND METHODS

Plant material: Alfalfa (*Medicago sativa* L. cv. Huaiyin) and (PL34HQ) were used as the recurrent male and non-recurrent female parent, respectively (Figure 1). A field trial was established in February 2020 at the grassland field to assess the multifoliate leaf populations at Yangzhou University, China.

Hybridization Development: Huaiyin alfalfa trifoliate leaves as the non-recurrent male parent and BC_1 with high multi-foliate leaves as the recurrent female parent were crossed to produce a BC_2 hybrid plant during 2018-2019. Seeds were harvested after maturity and sown in season 2020. Seeds of the BC_2 hybrids were sown in the field in PVC pipes 16cm in diameter and 50cm in height and placed concentrically on the soil surface (Figure 2). Pipes were placed 10 cm deep into the ground. The experimental unit was 26 meters long and 7 meters wide; all genotypes were planted in 7 rows; the distance between each row was 1 m, and the distance between the plants within the row was 1 meter (26 plants per row). In this study, 179 BC_2 populations (named BC_2-1 to BC_2-179) and their parents were used to evaluate genetic diversity among alfalfa multifoliate leaves.

Morphological Traits: During the study period, nine morphological traits were recorded in 179 BC_2 populations and two parents. These were; plant height (PH), number of branches per plant (NBP), stem diameter (SD), second branch length (SBL), number of multifoliate leaves per plant (NMFLP), number of multifoliate leaves on the main branch (NMFLMB), number of nodes (NN), fresh shoot weight (FSW), and dry shoot weight (DSW). Plant height was measured using a ruler (cm). A vernier caliper was used to measure the stem diameter 5 cm (mm) above the ground. The number of MF leaves per plant was calculated by dividing the number of branches with MF leaves by the total number of branches on that plant. The number of MF leaves on the main branch was determined by dividing the MF leaves on the branches by the total number of leaves in a single branch. To calculate the average value of MF leaves, five random branches from each plant were chosen. The fresh shoot weight (g) was calculated using a digital balance; the shoot dry weight (g) was calculated after oven-drying at 80°C for 48 hours until the plant reached a constant weight for biomass production.

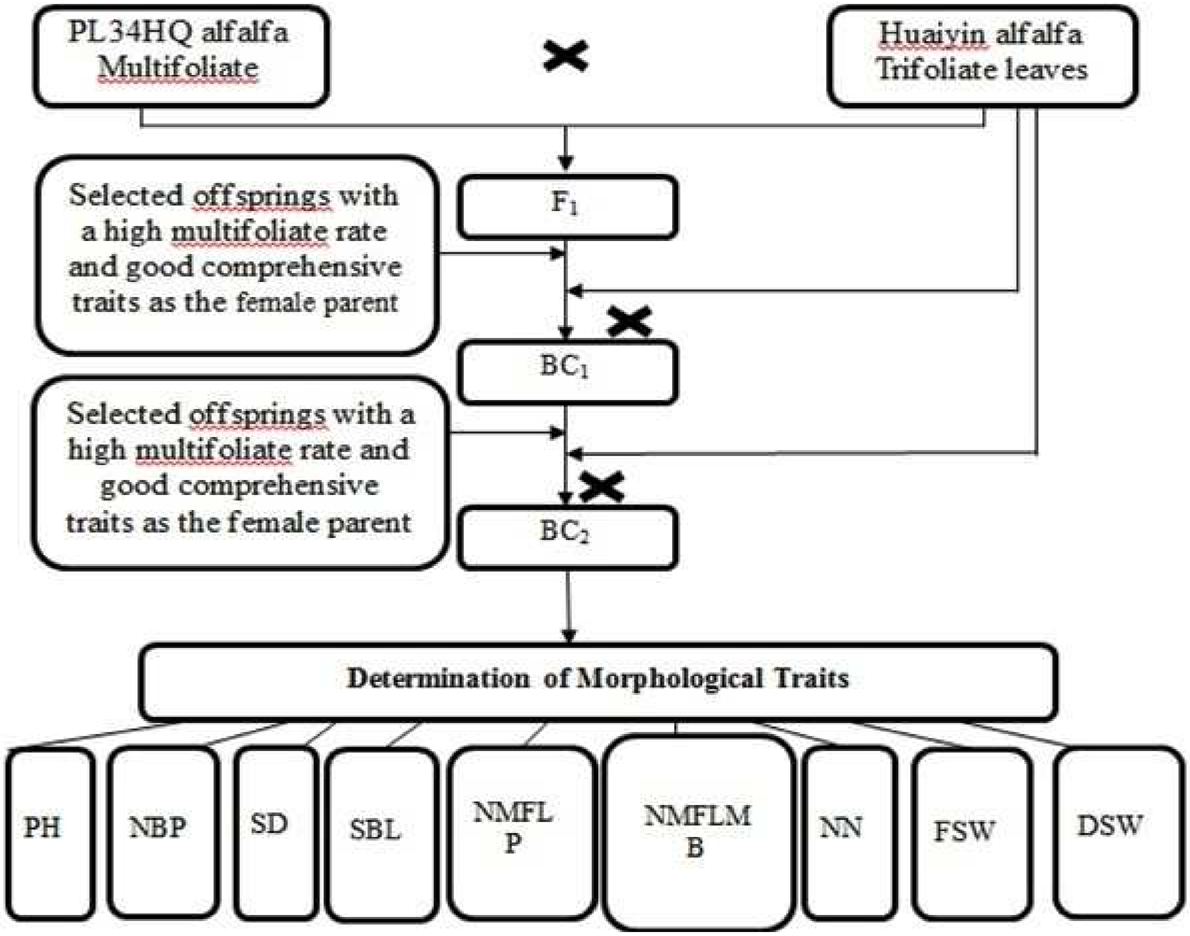

Fig 1. Hybridization scheme for backcrossing

Figure 2.BC$_2$populutions of alfalfa multifoliate leaves

Data analyses: The data recorded and then subjected to determine basic statistics (descriptive), and correlation coefficient (Anderson and Boyes, 1957).Data were further analyzed by principal component analysis (PCA) (Khodadadi *et al*., 2011) and cluster analysis (Everitt *et al*., 2001). The average hierarchical clustering method was used to construct the dendrogram based on morphological data. Data were performed with the R software version 4.1.0.

RESULTS

Statistical variation in the BC_2 population and two parents; PL34HQ and Huaiyin, is presented in Table 1. The range of the plant height of the BC_2 population of alfalfa was (26 – 98 cm) and the mean was 59.77 cm. Additionally, it was observed that 6.15% of the plant's height was higher than the high parent value, 29.05% was between two parents, and 64.80% was less than parental values. Moreover, genotype BC_2-72 recorded the greatest plant height (98.0 cm), while genotype BC_2-105 recorded the minimum plant height (26.0 cm). The range of the number of branches per plant of the 179 BC_2 population was (6.0- 74), and the average was (27.59). Furthermore, 42.45% of the number of branches per plant of the BC_2 population was higher than both parents, 19.55% of the number of branches per plant was between two parents, and 38% of the number of branches per plant was less than parental values. BC_2-105 had the lowest number of branches per plant (6.0), while BC_2-142 had the highest number of branches per plant (74). Our study showed that the minimum and maximum stem diameter of the 179 BC_2 population was 1.46 – 5.46 mm, with an average value of 2.76 mm. In addition, 3.91% of the BC_2 population had stem diameter as higher than the two-parent values, while 22.91% of the stem diameter was between the two-parent values, and 73.18% of the stem diameter was the lowest than the parental values. Moreover, genotype BC_2-19 showed a maximum stem diameter (5.46 mm), while the genotype BC_2-52 showed a minimum stem diameter (1.46mm). The range of the number of multifoliate leaves per plant of the179 BC_2 population was 0.00-90% and the mean value was 26.94%. The genotype BC_2-77 showed the highest number of multifoliate leaves per plant (90%), while the genotype BC_2-159 showed the lowest number of multifoliate leaves per plant, (1.8%). The number of multifoliate leaves on the main branch of the BC_2 population was distributed between 0.00 – 79.50%, and the mean value was12.29%. Furthermore, the range of the number of nodes of the179 BC_2 population was (12.00 - 24), and the average was (18.16). Among the population, 19% of the number of nodes was higher than the parental values, 77.65% was between parent values, and 3.35% was less than two parents. The minimum and maximum number of nodes (12.00 – 24.00) was observed in the genotype BC_2-155 and BC_2-40, respectively. Moreover, the minimum and maximum fresh shoot weight of the 179 BC_2 population were 3.10 – 359.4 g, and the mean value was 76.74 g. However, 30.73% of the fresh shoot weight of the BC_2 population was higher than both parents, 7.26% was between parents, and 62.01 was the lowest than the parental values. Lastly, the range of dry shoot weight of the BC_2 population was (0.40-108.60g), with an average value of 26.44 g. Among the population, 30.73% of the shoot dry weight was higher than the two-parent values, 7.26% was between two parents, and 62.01% was less than the parental values. The genotype BC_2-142 had the highest shoot fresh and dry weight (359.4 g – 108.6 g), whereas the genotype BC_2-92 showed the lowest shoot fresh and dry weight (3.10g – 0.40g).

Table 1.Statistical summary of nine morphological traits in two parentsandthe BC_2 populations.

Traits	Parents		179 BC2 population			
	Huaiyin	PL34HQ	Range	Minimum	Maximum	Mean ± Std. Error
PH	85.00	70.00	26.0-98.0	26.00	98.00	59.77 ± 1.26
NBP	27.00	23.00	6.0-74.0	6.00	74.00	27.59 ± 0.84
SD	4.31	3.34	1.46-5.46	1.46	5.46	2.76 ± 0.059
SBL	35.00	15.00	0.0-62.0	0.00	62.00	15.22 ± 1.02
NMFLP	0.00	95.70	0.0-90.0	0.00	90.00	26.94 ± 2.27
NMFLMB	0.00	90.60	0.00-79.5	0.00	79.50	12.29 ± 1.21
NN	20.00	13.00	12.0-24.0	12.00	24.00	18.16 ± 0.21
FSW	76.40	92.40	3.1-359.4	3.10	359.40	76.74 ± 5.83
DSW	24.20	32.10	0.4-108.6	0.40	108.60	26.44 ± 2.01

PH: plant height (cm), NBP: number of branches per plant, SD: stem diameter (mm), SBL: second branch length (cm), NMFLP: number of multifoliate leaves per plant, NMFLMB: number of multifoliate on the main branch, NN: number of nodes, FSW: fresh shoot weight (g), DSW: dry shoot weight (g).

Analysis of the correlation coefficient of the BC_2 population: Correlation analysis of nine agronomic traits of 179 genotypes of the BC_2 population and their parents is presented in Table 2. The result showed different degrees of correlation among plant height, number of branches per plant, stem diameter, second branch length, number of MF leaves per plant, number of MF on the main branch, number of nodes, fresh shoot and dry weight. Plant height was highly positively correlated with the stem diameter (r=0.787), followed by fresh shoot weight (r=0.669), dry shoot weight (r=0.668), second branch length (r=0.592), number of nodes (r=0.534), and number of branches per plant (r=0.323). There was a significant positive correlation between the number of branches per plant with fresh shoot weight (r=0.687), dry shoot weight (r=0.682), stem diameter (r=0.324), second branch length (r=0.225), and the number of nodes (r=0.183). Stem diameter and second branch length were in a significant relationship with dry shoot weight (r=0.727), fresh shoot weight (r=0.726), second branch length (r=0.641), number of nodes (r=0.409), and number of multifoliate leaves per plant (r=0.174). The number of MF leaves on the main branch and number of MF leaved per plant showed a positive correlation but not a statistically significant correlation with the plant height, number of branches per plant, nodes number, fresh shoot and dry weight. There was a highly significant positive correlation between the number of MF leaves per plant and the number of MF leaves on the main branch (r=0.673).

Principal component analysis (PCA): Principal component analysis (PCA) showed high genetic variation. The variation studied with principal components analysis showed that three main components with more than 1 eigenvalues higher accounted for 78.98% of the cumulative of the total variance among nine morphological traits (Table 3). The contribution rate of PC1 was the highest variability (48.8 %), followed by PC2 and PC3, which contributed 18.3% and 11.88% variability, respectively. Dry shoot weight, fresh shoot weight, stem diameter, plant height, second branch length, and node number were the main contributors to PC1. On the contrary, the main contributor to PC2 was the number of multifoliate leaves per plant and a number of multifoliate leaves on the main branch but contributed negatively. The number of branches per plant was a negative contribution to the PC3 (Table 4). The cumulative variance of 78.98% by the first three principal components with an eigenvalue of more than 1 indicated that the identified traits within the PCs significantly influenced the genotype and could be effectively used for selection. In this investigation, the first three eigenvalues correspond to the whole variance.

The correlation between groups was examined by plotting the first two principal components (PC1 x-axis and PC2 y-axis), which accounted for 67.1% of the total variance (Fig.4). It can be seen that the number multifoliate leaves on the main branch was significantly positively correlated with the number of multifoliate leaves per plant. In contrast, branch number was positively correlated with plant height, stem diameter, second branch length, number of nodes, shoot fresh weight and shoot dry weight.

Table 2.Correlation coefficient among nine morphological traits in BC_2 populations.

Traits	PH	NBP	SD	SBL	NMFLP	NMFLMB	NN	FSW	DSW
PH	1								
NBP	0.323^{**}	1							
SD	0.787^{**}	0.324^{**}	1						
SBL	0.592^{**}	0.225^{**}	0.641^{**}	1					
NMFLP	0.120	0.002	0.174^{*}	0.167^{*}	1				
NMFLMB	0.020	0.060	0.041	0.097	0.673^{**}	1			
NN	0.534^{**}	0.183^{*}	0.409^{**}	0.271^{**}	0.108	0.010	1		
FSW	0.669^{**}	0.687^{**}	0.726^{**}	0.518^{**}	0.107	0.016	0.352^{**}	1	
DSW	0.668^{**}	0.682^{**}	0.727^{**}	0.528^{**}	0.112	0.028	0.358^{**}	0.983^{**}	1

*, Significant at $P< 0.05$; **, Significant at $P< 0.01$, PH: plant height (cm), NBP: number of branches per plant, SD: stem diameter (mm), SBL: second branch length (cm), NMFLP: number of multifoliate leaves per plant, NMFLMB: number of multifoliate leaves on the main branch, NN: number of nodes, FSW: fresh shoot weight (g), DSW: dry shoot weight (g).

Table 3. Principal component analysis eigenvalues, variability rate percentage and cumulative values.

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Eigenvalues	4.36	1.66	1.07	0.76	0.41	0.32	0.23	0.18	0.02
Variability rate %	48.8	18.3	11.88	8.24	4.64	3.46	2.54	1.95	0.18
Cumulative rate %	48.8	67.1	78.98	87.23	91.86	95.32	97.86	99.81	100.00

Table 4.The principal component analysis among nine morphological traits of the BC_2 populations.

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
PH	0.844	0.039	0.327	0.006	0.123	0.240	0.266	0.190	0.002
NBP	0.617	0.146	-0.661	0.208	-0.225	0.007	0.229	-0.120	0.001
SD	0.865	-0.001	0.234	-0.190	0.220	0.094	-0.025	-0.321	0.000
SBL	0.698	- 0.088	0.286	-0.463	-0.436	-0.138	-0.006	0.024	0.002
NMFLP	0.207	- 0.891	-0.015	0.034	0.185	-0.325	0.143	0.009	0.001
NMFLMB	0.102	-0.899	-0.176	0.045	-0.143	0.331	-0.137	0.005	0.002
NN	0.531	0.002	0.464	0.679	-0.164	-0.080	-0.087	-0.036	0.001
FSW	0.921	0.122	- 0.272	-0.010	0.121	-0.063	-0.161	0.103	0.093
DSW	0.923	0.111	- 0.265	-0.011	0.109	-0.063	-0.175	0.105	0.092

PH: plant height (cm), NBP: number of branches per plant, SD: stem diameter (mm), SBL: second branch length (cm), NMFLP: number of multifoliate leaves per plant, NMFLMB: number of multifoliate leaves on the main branch, NN: number of nodes, FSW: fresh shoot weight (g), DSW: dry shoot weight (g).

Cluster analysis: Clustering analysis was used to determine genetically related genotypes and which genotypes would be suitable for future breeding programs. Based on nine agronomic traits for 179 genotypes and their parents, cluster analysis grouped them into four discrete and well-defined clusters (Table 5 and Fig 5). Cluster II included 135 different genotypes, followed by 40, 4 and 2 genotypes in clusters III, IV, and I, respectively. The variability of 179 genotypes and their parents was also reflected in the cluster means for the nine traits evaluated (Table 6). Cluster I had two genotypes constituting1.1% of the total population. They were both characterized by genotypes with a high number of multifoliate leaves per plant (92.85%) and number of multifoliate leaves on the main branch (85.05%). Cluster II had 135 genotypes, which made up 74.59% of the total population. Genotypes in Cluster II were distinguished by their low plant height (54.67 cm), the number of branches per plant (23.58), stem diameter (2.48 mm), second branch length (11.61 cm), multifoliate leaves per plant (23.87%), fresh shoot weight (37.39 g), and dry shoot weight (12.91 g). Cluster III comprised 40 genotypes, accounting for 22.1% of the total population with high plant height (76.1 cm). Cluster IV had four genotypes that made up 2.21% of the total population, characterized by increased number of branches per plant (45), stem diameter (3.89mm), second branch length (30 cm), number of nodes (20.5), fresh shoot weight (314.15 g) and dry shoot weight (107.08 g).

Table 5.Distribution of the two parents and the BC_2 populations of alfalfa multifoliate leaves in various clusters.

Cluster	No. of genotypes	Name of genotypes
I	2	BC_2-77, non-recurrent parent (PL34HQ)
II	135	BC_2-41, BC_2-46, BC_2-15, BC_2-94, BC_2-69, BC_2-77, BC_2-14, BC_2-30, BC_2-72, BC_2-22, BC_2-98, BC_2-101, BC_2-31, BC_2-127, BC_2-59, BC_2-11, BC_2-58, BC_2-42, BC_2-53, BC_2-44, recurrent parent (Hyaiyin), BC_2-21, BC_2-122, BC_2-2, BC_2-34, BC_2-13, BC_2-125, BC_2-12, BC_2-129, BC_2-169, BC_2-104, BC_2-57, BC_2-111, BC_2-105, BC_2-131, BC_2-164, BC_2-170, BC_2-87, BC_2-166, BC_2-146, BC_2-148, BC_2-147, BC_2-27, BC_2-133, BC_2-137, BC_2-174, BC_2-151, BC_2-167, BC_2-89, BC_2-149, BC_2-17, BC_2-79, BC_2-150, BC_2-172, BC_2-54, BC_2-56, BC_2-66, BC_2-24, BC_2-28, BC_2-102, BC_2-135, BC_2-154, BC_2-155, BC_2-136, BC_2-9, BC_2-10, BC_2-138, BC_2-156, BC_2-25, BC_2-65, BC_2-50, BC_2-82, BC_2-168, BC_2-84, BC_2-132, BC_2-51, BC_2-134, BC_2-119, BC_2-83, BC_2-86, BC_2 -80, BC_2-81, BC_2-113, BC_2-153, BC_2-158, BC_2-157, BC_2-71, BC_2-1, BC_2-118, BC_2-156, BC_2-160, BC_2-173, BC_2-130, BC_2-36, BC_2-61, BC_2-23, BC_2-107, BC_2-109, BC_2-117, BC_2-38, BC_2-37, BC_2-55, BC_2-63, BC_2-152, BC_2-93, BC_2-106, BC_2-52, BC_2-110, BC_2-29, BC_2-178, BC_2-67, BC_2-78, BC_2-39, BC_2-108, BC_2-62, BC_2-26, BC_2-163, BC_2-35, BC_2-92, BC_2-90, BC_2-112, BC_2-85, BC_2-88, BC_2-91, BC_2-175, BC_2-64, BC_2-116, BC_2-115, BC_2-114, BC_2-123, BC_2-45, BC_2-179, BC_2-60, BC_2-40, BC_2-176.
III	40	BC_2-95, BC_2-18, BC_2-161, BC_2-141, BC_2-8, BC_2-48, BC_2-75, BC_2-74, BC_2-76, BC_2-143, BC_2-139, BC_2-7, BC_2-16, BC_2-4, BC_2-47, BC_2-70, BC_2-3, BC_2-6, BC_2-19, BC_2-121, BC_2-140, BC_2-177, BC_2-159, BC_2-33, BC_2 -144, BC_2-43, BC_2-100, BC_2-99, BC_2-96, BC_2-17, BC_2-49, BC_2-5,

		BC_2-3, BC_2-128, BC_2-145, BC_2-68, BC_2-73, BC_2-120, BC_2-103, BC_2-124.
IV	4	BC_2-142, BC_2-162, BC_2-20, BC_2-126.

Table 6. Mean values of nine morphologicaltraits of the BC_2 population in cluster analysis.

Custer	PH	NBP	SD	SBL	NMFLP	NMFLMB	NN	FSW	DSW
I	67.5	31.5	3.28	23	92.85	85.05	16	86.75	29.75
II	54.67	23.58	2.48	11.61	23.87	11.2	17.7	37.39	12.91
III	76.1	39.05	3.59	26.03	35.29	14.86	19.4	185	63.95
IV	74	45	3.89	30	24.32	3.75	20.5	314.15	107.1

PH: plant height (cm), NBP: number of branches per plant, SD: stem diameter (mm), SBL: second branch length (cm), NMFLP: number of multifoliate leaves per plant, NMFLMB: number of multifoliate leaves on the main branch, NN: number of nodes, FSW: fresh shoot weight (g), DSW: dry shoot weight (g).

Biplot (axes PC1 and PC2: 67.1)

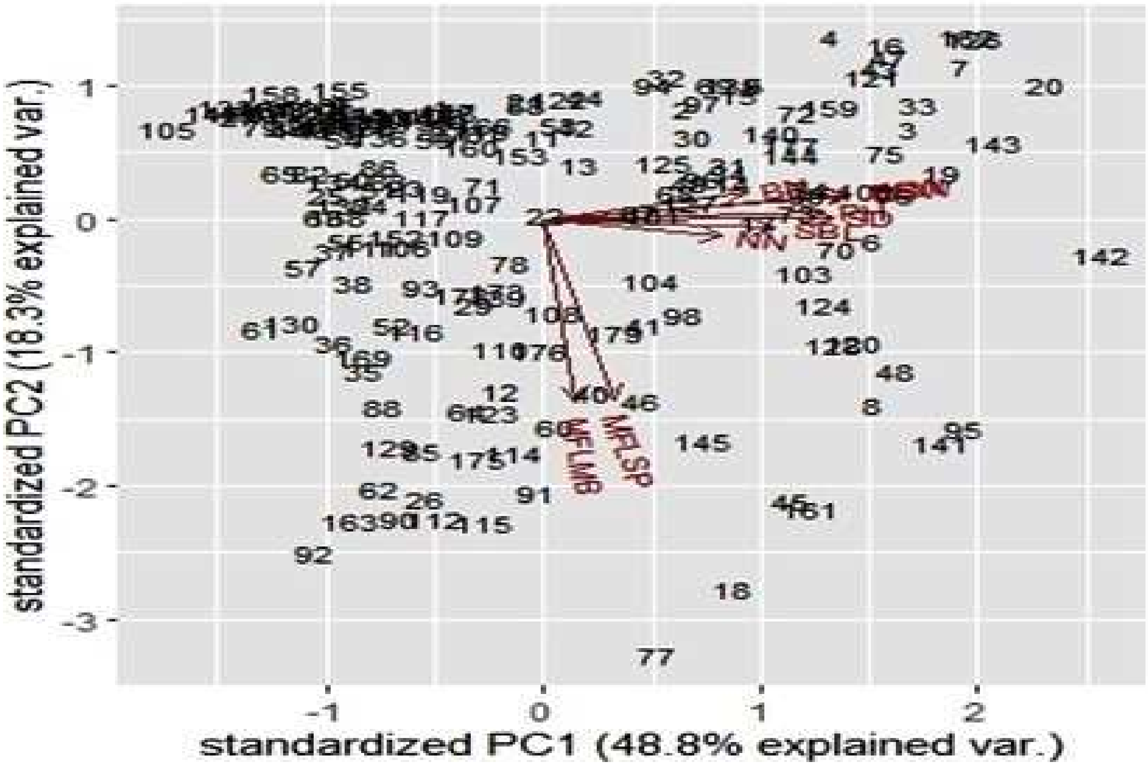

Fig 3.Biplot of morphological traits of the two parents and the BC_2 populations for principal components PC1 and PC2.

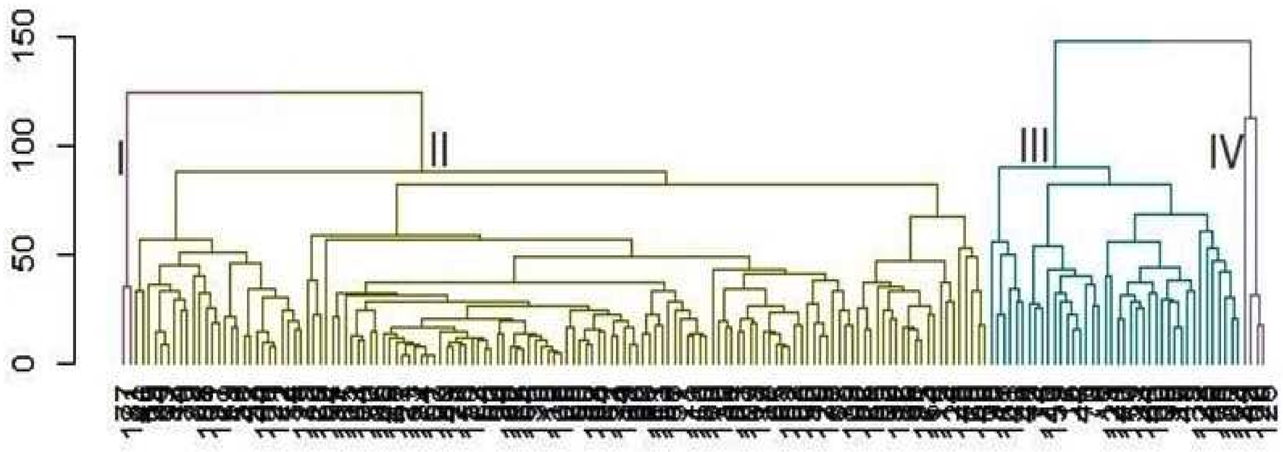

Fig 5.Cluster dendrogram analysis of nine morphological traits of the two parents and the BC_2 populations.

DISCUSSION

Alfalfa is considered "the queen of forage crops" because of it is biomass, biological nitrogen fixation, different nutritional profiles, stress resistance, and consistent protein and mineral supplies for cattle. Backcross breeding was done between the PL34HQ (multifoliate leaves) and *Medicago sativa* L. cv. Huaiyin (trifoliate leaves). It is possible to transfer the high yield and good genes of Huaiyin alfalfa into the genetic background of alfalfa MF leaves with no change in its leaflet characteristics of Huaiyin intact. The combination of high yield and multifoliate traits of PL34HQ creates a high-yielding, and high-quality new alfalfa material with MF characteristics.

In this study, Huaiyin alfalfa was used as the recurrent parent, and alfalfa PL34HQ multifoliate was used as the non-recurrent parent. At the pod setting stage, we obtained the BC_2 population,which hada 66.85% multifoliate rate. Nine morphological traits of the BC_2 population were studied to select genotypes based on the best morphological traits for future breeding programs through descriptive statistics, correlation coefficient, principal component analysis, and cluster analysis. The fresh shoot weight, number of multifoliate leaves per plant, dry shoot weight, number of multifoliate leaves on the main branch, plant height, second branch length, and the number of branches per plant showed high variation.

The alfalfa yield was determined by the plant height, the number of branches per plant, and shoot fresh and dry weight. The correlation coefficient is a statistical measurement used to evaluate the relationship between the examined morphological traits. The current study showed a strong positive correlation between plant height and number of branches per plant, stem diameter, fresh shoot weight, dry shoot weight, second branch length, and the number of nodes. It means that an increase in one trait would increase another. The study showed that the shoot fresh weight had a high positive correlation with the dry shoot weight. On the contrary, a negative correlation of shoot fresh weight with dry matter was reported by (Munir *et al*., 2021).This result may indicate that the different populations have different morphological traits. Additionally, the environment may impact the differences between these traits. Yield-related traits such as branch number, second branch length, node number, and stem diameter will increase accordingly, effectively improving yield.

Meanwhile, alfalfa stem diameter and number of nodes will increase with alfalfa plant height. Stem diameter and second branch length positively correlated with fresh shoot weight, dry shoot weight, number of nodes and number of multifoliate leaves per plant. Number of branches per plant positively correlated with fresh shoot weight, dry shoot weight, stem diameter, plant height, second branch length and number of nodes. Similar findings were reported by Petkova (2010), who observed positive correlation of green mass yield with plant height and the number of stems in MF 23 plants. However, the number of MF leaves on the main branch showed no significant correlation with all morphological traits except the number of MF leaves per plant. Moreover, our results showed that the MF traits could perform better in the BC_2 population. The MF traits of the backcross population showed a specific positive correlation with the yield-related traits. Still, it was not significant.

The principal component analysis is an effective tool for obtaining parental lines for successful breeding programs (Amna *et al*., 2013). The principal component analysis is widely used to evaluate genotypes based on morphological traits and their grouping (Rizwana *et al*., 2010; Sajjad *et al*., 2011; Janmohammadi *et al*., 2014). Eigenvalue helps select principal components that have the highest contribution in variation. Variation is one of the essential requirements for the selection phase in plant breeding (Canci and Toker, 2014; Benlioglu and Ozkan, 2022). The first three components had an eigenvalue above 1.0. These three main components accounted for 78.98% of the total variation. The first principal component is more related to yield traits Benlioglu and Ozkan (2022) reported that the principal components with an eigenvalue greater than 1 accounted for 80.57% of the total variation and that the first principal component was more related to the biomass per plot and plant height. Also, Saleem *et al*. (2023) observed that the first two main components made up 80.82% of the total variation. The first principal components in this study included dry shoot weight, fresh shoot weight, stem diameter, plant height, second branch length, and number of branches per plant. Similar findings were reported by Saleem *et al*. (2023), who noted that the first principal component was more closely related to plant height, leaf number of plant, and stem diameter. Another study revealed 62.83% of genotype variability in the first two principal components. The PC1 showed positive factor loadings for all characteristics exclusively related to yield and yield-related components (Khalid *et al*., 2022). The scree plot explains the percentage variation linked with each of the components. The eigenvalue significance criterion (Kaiser, 1960) was used to select statistically significant principal components. Based on the principal component analysis, genotypes with the highest score and desirable characters can be selected for further breeding programs.

Cluster analysis could be considered an effective tool to classify germplasm, which provides a consistent foundation for selecting base material to design breeding tactics in the future (Sušić *et al*., 1999; Jin *et al*., 2006; Mohan *et al*., 2021; Benlioglu and Ozkan, 2022). Cluster analysis grouped the 179 BC_2 population of alfalfa multifoliate leaves and their parents into four clusters with some distinct traits. The heterotic effect in cross progenies can be increased by grouping the lines into separate, distinct clusters. The lines with desired traits can used in crossbreeding programs to obtain higher hybrid vigor and improved segregants. The genotypes grouped under clusters I and IV had the highest morphological traits, such as NMFLP and NMFLMB in cluster I, and NBP, SD, SBL, NN, FSW, and DSW in cluster IV. They can be utilized as potential donors in the future crossing program for improving high multifoliate leaves and yield traits.

Conclusion: Plants get their genetic diversity by picking and backcrossing parents with desirable traits. This study showed a significant and wide variation in most morphological traits studied in BC_2 populations, which could be used to improve hybrid breeding programs. In principal component analysis, it was shown that three principal components (PC1, PC2, and PC3) accounted for 78.98% of the total variation. It was an excellent method to find out which traits differed most between genotypes and which variables to use to group genotypes. The cluster analysis method grouped the genotype into four main clusters based on different morphological traits. Cluster I and IV have genotypes with highest morphological traits, such as BC_2-77 and non-recurrent parent (PL34HQ), which had the highest NMFLP and NMFLMB in cluster I, and (BC_2-142, BC_2-162, BC_2-20, BC_2-126), which had the highest NBP, SD, SBL, NN, FSW, and DSW in cluster IV. It was suggested that these genotypes can be chosen for hybridization to establish future hybrids having multifoliate characteristics.

Acknowledgements: This study was funded by the Jiangsu Students Innovation and Entrepreneurship Training Program (grand number: 202211117098Y).

REFERENCES

Adhikari, L., O. M. Lindstrom, J. Markham and A. M. Missaoui (2018). Dissecting key adaptation traits in the polyploid perennial Medicago sativa using GBS-SNP mapping. Frontiers Plant Sci.,9: 934. https://doi.org/10.3389/fpls.2018.00934.

Amna, N., F. Jehanzeb, M. Abid, S. Muhammad and R. Muhammad (2013). Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of *Gossypium hirsutum* L. Вестник аграрной науки. 43(4): 2-9.

Anderson, E. and J. W. Boyes (1957). An experimental investigation of judgements concerning genera and species. Evolution. 260-263. https://doi.org/10.2307/2406055.

Annicchiarico, P., C. Scotti, M. Carelli and L. Pecetti (2010). Questions and avenues for lucerne improvement. Czech J Genet. Plant Breed., 46(1): 1-13. https://doi.org/10.17221/90/2009-CJGPB.

Benlioglu, B. and U. Ozkan (2022). Multivariate analysis and its application for screening mungbean [*Vigna radiata* (L.) wilczek] landraces. Legume Research-An International Journal. 1: 6. DOI : 10.18805/LRF-661.

Biazzi, E., N. Nazzicari, L. Pecetti, E. C. Brummer, A. Palmonari, A. Tava and P. Annicchiarico (2017). Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. PLoS One. 12(1): e0169234.https://doi.org/10.1371/journal.pone.0169234.

Bingham, E. and R. Murphy (1965). Breeding and Morphological Studies on Multifoliolate Selections of Alfalfa, *Medicago sativa* L. 1. Crop Science. 5(3): 233-235. https://doi.org/10.2135/cropsci1965.0011183X000500030010x.

Brummer, E. C. (1999). Capturing heterosis in forage crop cultivar development. Crop science. 39(4):

943-954.https://doi.org/10.2135/cropsci1999.0011183X003900040001x.

Canci, H. and C. Toker (2014). Yield components in mung bean (*Vigna radiata* (L.) Wilczek). Turkish Journal of Field Crops. 19(2): 258-261. https://doi.org/10.17557/tjfc.79162.

Dangi, R., A. Kumar and A. Khar (2018). Genetic variability, heritability, and diversity analysis studies in short day tropical onion (*Allium cepa* L.). Ind. J. Agric. Sci. 88: 948-957.

Everitt, B., S. Landau and M. Leese (2001). Cluster Analysis. A Hodder Arnold Publication. Willey, London.

Janmohammadi, M., Z. Movahedi and N. Sabaghnia (2014). Multivariate statistical analysis of some traits of bread wheat for breeding under rainfed conditions. Journal of Agricultural Sciences (Belgrade). 59(1): 1-14.

Jin, F., J. Xue, Y. Jia and Z. Liu (2006). The cluster analysis on tomato germplasms. Acta Agri. Boreali-Sin. 21: 49-54.

Juan, N. A., C. C. Sheaffer, D. K. Barnes, D. R. Swanson and J. H. Halgerson (1993). Leaf and stem traits and herbage quality of multifoliolate alfalfa. Agronomy Journal. 85(6): 1121-1127. https://doi.org/10.2134/agronj1993.00021962008500060005x.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and psychological measurement. 20(1): 141-151.https://doi.org/10.1177/001316446002000116.

Khalid, A. and A. Hammed (2021). Genetic divergence in wheat genotypes based on seed biochemical profiles appraised through agglomerative hierarchical clustering and association analysis among traits. Pakistan J. Bot. 53(4): 1281-1286.DOI: http://dx.doi.org/10.30848/PJB2021-4(7).

Khalid, A., A. Hameed and M. F. Tahir (2022). Estimation of genetic divergence in wheat genotypes based on agro-morphological traits through agglomerative hierarchical clustering and principal component analysis. Cereal Research Communications. (51): 217–224.https://doi.org/10.1007/s42976-022-00287-w.

Khan, A., M. Ahmad, M. K. N. Shah and M. Ahmed (2020). Performance of wheat genotypes for Morpho-Physiological traits using multivariate analysis under terminal heat stress. Pakistan J. Bot. 52(6): 1981-1988. DOI: http://dx.doi.org/10.30848/PJB2020-6(30).

Khodadadi, M., M.H. Fotokian and M, Miransari (2011). Genetic diversity of wheat (*Triticum aestivum* L.) genotypes based on cluster and principal component analyses for breeding strategies. Aust. J. Crop. Sci. 5: 17-24.

Li, X. and E. C. Brummer (2009). Inbreeding depression for fertility and biomass in advanced generations of inter-and intrasubspecific hybrids of tetraploid alfalfa. Crop science. 49(1): 13-19. https://doi.org/10.2135/cropsci2008.04.0205.

Li, X. and E. C. Brummer (2012). Applied genetics and genomics in alfalfa breeding. A gronomy J. 2(1): 40-61. https://doi.org/10.3390/agronomy2010040.

Mishra, C., V. Tiwari, V. G. Satish-Kumar, A. Kumar and I. Sharma (2015). Genetic diversity and genotype by trait analysis for agromorphological and physiological traits of wheat (*Triticum aestivum* L.). Sabrao J. Breed. Genet. 47(1): 40-48.

Mohan, S., A. Sheeba and T. Kalaimagal (2021). Genetic diversity and association studies in greengram (*Vigna Radiata* (L.) Wilczek). Legume Research-An International Journal. 44(7): 779-784.DOI : 10.18805/lr-4176.

Munir, M. A., M. Z. Mushtaq and A. Ghani (2021). Genetic Diversity Estimation in Alfalfa (*Medicago sativa* L.) Genotypes using SSR Markers, Morphological and Yield Related Traits. J. Agric. Res. 59(2): 133-139.

Petkova, D. (2010). Multifoliate Alfalfa line with 23-24 leaves on a leaf stalk. Journal of Crop and Weed. 6(1): 1-5.

Riday, H. and E. C. Brummer (2002). Forage yield heterosis in alfalfa. Crop Science. 42(3): 716-723.https://doi.org/10.2135/cropsci2002.7160.

Rizwana, M., S. Muhammad, K. Ihsan, A. Khan and S. Khan (2010). Morphological diversity and traits association in bread wheat (*Triticum aestivum* L.). American-Eurasian Journal of Agricultural and Environmental Science. 8(2): 216-224.

Sajjad, M., S. Khan and A. S. Khan (2011). Exploitation of germplasm for grain yield improvement in spring wheat (*Triticum aestivum*). Int. J. Agric. Biol. 13(5): 695-700.

Shi, S., L. Nan and K. F. Smith (2017). The current status, problems, and prospects of alfalfa (*Medicago sativa* L.) breeding in China. Agronomy. 7(1):1. https://doi.org/10.3390/agronomy7010001.

Saleem, A. M., K. Ziaf, M. Amjad, A. Shakeel, M. A. Ghani and A. Noor (2023). Assessment of genetic diversity among okra genotypes through PCA and correlation analysis for fruit tenderness, and morphological and yield traits. Pakistan J. Bot. 55(2): 555-562.DOI: http://dx.doi.org/10.30848/PJB2023-2(34).

Sušić, Z., J. Zdravković, N. Pavlović and S. Prodanović (1999). Selecting features for estimating genetic divergence of tomato genotypes (*Lycopersicum esculentum* Mill.). Genetika. 31(3): 235-244.