GMR
The Original

Evaluation of Genetic Diversity and Population Structure in Viola Species by the Application of the SRAP Marker System

S. Latif[1], A. R, Khan[1], A. Nazir[2], R. Ahmad[1] and S. M. Shah[1]*

1) Department of Biotechnology, COMSATS University Islamabad, Abbottabad Campus, Abbottabad, Pakistan
2) Department of Environmental Sciences, COMSATS University Islamabad, Abbottabad Campus, Abbottabad, Pakistan

*Corresponding author: S.M. Shah, E-mail: smasood@cuiatd.edu.pk

ABSTRACT

Viola species are perennial herbs belonging to the family Violaceae. It is essential to evaluate genetic diversity due to its applications in the food and pharmaceutical industries, as this is essential for genetic improvement. The present study was designed to explore genetic diversity in seven Viola species collected from twelve locations of KP, Pakistan. A set of 42 sequence-related amplified polymorphism markers was used to study the genetic diversity of these species. A total of 456 bands were amplified, and the detected bands were 100% polymorphic. The highest PIC value (0.32) was obtained for the Me2Em7 and Me3xEm6 primer pairs, and the primer pair Me4Em4 had the lowest PIC value of 0.17, with a mean PIC value of 0.25. The major allele frequency ranged from 0.68 to 0.86, with an average of 0.79. The minor allele frequency ranged from 0.14 to 0.32, with a mean of 0.21. The average gene diversity was 0.30, ranging from 0.21 (Me4Em4) to 0.41 (Me2Em7 and Me3Em6). The mean value of Shannon's index was 1.88, varying from 0.03 to 2.81. The highest Shannon's index value was obtained for the Me1Em4 primer pair. Five distinct clusters were identified through neighbor-joining analysis. The *Viola hirta* collected from Mushkpuri exhibited the greatest genetic divergence from the areas studied. The structure results pointed out the influence of geographic and environmental variations on genetic diversity within the same species, along with intra-species genetic conservation at a certain level as well. PCoA analysis reveals three clusters are in the sport of results obtained from NJ and structure. In a nutshell, divergence patterns and distinct clustering patterns among the studies' genotypes highlighted valuable genetic resources and their application as food and pharmaceuticals. For ensuring the sustainable supply of this valuable medicinal plant and developing effective conservation strategies, these findings provide a foundation.

Key words: *Medicinal plant; Genetic variation; Population structure*

INTRODUCTION

The diverse family, Violaceae comprising 1100 species of trees, shrubs, lianas, and herbs (Wahlert et al. 2018). The Violaceae family has 26 genera that vary among fruit type, growth, shape, inflorescence configurations, and floral morphology (Flicker and Ballard, 2015). The best-known genus, which comprises over 500 species, is Viola, which is distributed globally in temperate regions of the tropical high-altitude locations and the Northern Hemisphere. On stalks, they produced a combination of flowers that could be white, violet, yellow, blue and purple in colour. This diversity is the essence of the biological world in changing circumstances and is crucial to the survival of species as well. In medicinal plants, genetic diversity is also a vital factor in their potential applications and medicinal properties (El-Shazly et al. 2020). Medicinal plants that are genetically diverse are more resilient to environmental stress and diseases. This property makes them more sustainable for commercial application and cultivation.

DNA-based diversity assessments have proved to be long-lasting and have diverse applications. Several molecular marker systems, including SNP, AFLP, SSR, RFLP, SRAP and RAPD, are used for genetic characterization of a population (Nam et al. 2020, Abolghasemi et al. 2020). SRAP molecular markers have multi-allelic characteristics, non-specific, target

coding regions, co-dominant and cost-effective. These characteristics make them suitable and more effective for genetic diversity studies (Li and Quiros, 2001; Robarts and Wolfe, 2014). To determine genetic diversity in crop plants PCR-based SRAP markers were applied by number of researchers around the globe since their development (Li and Quiros 2001; Aneja et al. 2012). In addition to crops these markers are used to assess medicinal plants genetic diversity such as *Salvia miltiorrhiza*, Pinellia, *Mallotus oblongifolius*, *Chrysanthemum morifolium,* and *Nigella sativa* (Liu et al. 2012; Golkar and Nourbakhsh, 2019).

Pakistan, the land of diverse environments, has a rich plant diversity. The major portion of this diversity is native to the northern areas. Viola species are part of this natural resource as well. They have been used in herbal medicine for centuries to cure several diseases (Wang et al. 2025; Batiha et al. 2023). There is an urge to explore phytochemical diversity, and hence several researchers focused on this property for investigation, including Pakistan (Fatima et al. 2023; Dastagir et al. 2023). Production of these phytochemicals is controlled by many factors, not limited to genetics, but this factor is thought to be the key player.

Therefore, the importance of genetic variation and its influence on the medicinal properties of Viola species was reviewed. The genetic diversity of Viola species at the molecular level in northern Pakistan is least studied. Therefore, this study aimed to provide the first documentation of various aspects related to the genetic diversity assessment of this genus of medicinal plants. Viola species diversity was examined using a SRAP molecular framework. Additionally, identifying polymorphic markers and evaluating genetic variation is crucial for developing suitable breeding plans and for conserving Viola species in the region.

MATERIALS AND METHODS

Collection and Identification of Viola species

Seven species of Viola were collected from twelve locations of KP, Pakistan (Table 1). For each genotype, five samples were collected from each site. Voucher specimens were collected for each genotype and deposited in the herbarium of the Department of Environmental Sciences, COMSATS University Islamabad, Abbottabad Campus, Abbottabad, Pakistan. The collected plant specimens were identified by an expert taxonomist and verified using the Flora of Pakistan. To ensure nomenclatural accuracy, the botanical names and corresponding family classifications were further validated through the international database "The World Flora online "(https://www.worldfloraonline.org/).

Table 1. *Viola* species name, voucher No. and the collection site

S. No.	Species/ Voucher Number	Locations	Codes	Coordinates	Altitude (m)
1	*Viola mirabilis* L. CUHA-477	Shogran	*V. mirabilis_ Sh*	34.6396 N, 73.4689 E	2,390
2	*Viola pilosa* Blume CUHA-108	Ayubia	*V. pilosa_ Ay*	34.0754 N, 73.4007 E	2,400
		Namlimera	*V. pilosa_ Nm*	34.0765 N, 73.3931 E	2,410
		Chattar Plain	*V. pilosa_ C*	34.1833 N, 73.3086 E	1,800
		Thandiani	*V. pilosa_ T*	34.2314 N, 73.4141 E	2,625
		Baragali	*V. pilosa_ B*	34.0792 N, 73.3810 E	2,300
		Miandam	*V. pilosa_ M*	35.0061 N, 72.5530 E	1,320
3	*Viola hirta* L. CUHA-478	Mushkpuri	*V. hirta_ Mu*	34.0727 N, 73.4012 E	2,800
4	*Viola canescens* Wall. CUHA-479	Abbottabad	*V. canescens_ A*	34.1688 N, 73.2215 E	1,260
5	*Viola odorata* L. CUHA-480	Abbottabad	*V. odorata_ A*	34.1688 N, 73.2215 E	1,260
		Bahrain	*V. odorata_ Ba*	35.0038 N, 72.8996 E	1,430
		Shogran	*V. odorata_ Sh*	34.6396 N, 73.4689 E	2,390
		Nathiagali	*V. odorata_ N*	34.0617 N, 73.3903 E	2,410
		Thandiani	*V. odorata_ T*	34.2314 N, 73.4141 E	2,625
6	*Viola biflora* L. CUHA-481	Swat	*V. biflora_ S*	35.2184 N, 72.4258 E	2,500
		Shogran	*V. biflora_ Sh*	34.6396 N, 73.4689 E	2,390
7	*Viola cucullate* Aiton CUHA-482	Shogran	*V. cucullata_ Sh*	34.6396 N, 73.4689 E	2,390

Isolation of Total Genomic DNA

Young leaves were collected, packed in a zipper bag, labeled, and stored at -20°C till used for DNA extraction. Total genomic DNA was extracted from all genotypes individually by using the CTAB method described by Doyle (1991) with some modifications. Briefly, leaves were ground to a fine paste by adding 600 µl Lysis Buffer 2% CTAB (1M Tris-HCl, pH 8.00, 0.25M EDTA, 4g CTAB, and 16 g NaCl) and kept in a water bath at 65ºC for 15 min, followed by centrifugation (8000 rpm for 5 min) at 4ºC. Then, 600 µl of Chloroform: iso-amyl alcohol (24:1, v/v) was added to the supernatant. Then the same volume of chilled isopropanol was added to a fresh tube containing supernatant, left overnight at -20ºC, and centrifuged for 5 min at 8000 rpm at 4ºC. The collected pellets were washed with chilled 70% (v/v) ethanol, air-dried, re-mixed in TE buffer, and kept at -20°C for further use. The DNA extracted was quantified by Gel Electrophoresis using 1% agarose with 5x TBE. For PCR, the concentration was normalized to 50 ng/µL of DNA.

SRAP-PCR conditions

The polymerase chain reaction was performed in a volume of 15 µL containing 5 µL double-distilled water, 1 µL of each forward and reverse primer, 5 µL of 2X Master Mix (by Zokeyo), and 3 µL of DNA. The thermal cycling profile was 95 °C for 5 minutes, followed by 5 cycles of 94 °C for 30 sec, 37 °C for 30 sec, 72 °C for 30 sec, 35 cycles of 94 °C for 30 sec, 47 °C for 30 sec, and 72 °C for 30 sec. The thermal cycling profile was the same for all combinations except the annealing temperature, which ranged from 35 °C to 47 °C (Supplementary Table 1). All PCR reactions were performed in a thermocycler (by Kyratec). Gel electrophoresis was used to separate amplified products in 2% agarose gel in 5x TBE buffer at a constant voltage (120 V) for 25-30 minutes. The gel was treated with ethidium bromide (10 mg/ ml) for UV visibility.

Forty-two SRAP primer combinations were used for screening, and forward and reverse sequences of SRAP primers are provided in Table 2. All primer combinations produced amplified bands under the optimized conditions. The amplifications were repeated twice, and only the visible repetitive bands were used in data analysis.

Table 2. SRAP primers with their sequences used for the current study

S no.	Forward Primers	Sequence 5'-3' (17 bp)	Reverse Primers	Sequence 5'-3' (18 bp)
1	Me1	TGAGTCCAAACCGGATA	Em1	GACTGCGTACGAATTAAT
2	Me2	TGAGTCCAAACCGGAGC	Em2	GACTGCGTACGAATTTGC
3	Me3	TGAGTCCAAACCGGAAT	Em3	GACTGCGTACGAATTGAC
4	Me4	TGAGTCCAAACCGGACC	Em4	GACTGCGTACGAATTTGA
5	Me5	TGAGTCCAAACCGGAAG	Em5	GACTGCGTACGAATTAAC
6	Me6	TGAGTCCAAACCGGACA	Em6	GACTGCGTACGAATTGCA
7			Em7	GACTGCGTACGAATTCAA

Data Scoring and Statistical Analysis

Amplified fragments with the same gel mobility were scored as 1 for presence or 0 for absence to generate a binary data matrix. The bands that are common in all genotypes were referred to as monomorphic, and those that differed among the genotypes were referred to as polymorphic, irrespective of their band strength. Power Marker (V3.25) (Liu and Muse, 2005) was used to find genetic diversity among all Viola genotypes by analyzing four parameters: polymorphism information content, major allele frequency, minor allele frequency, and gene diversity. Shannon's index was calculated using Equation 1

$$H=-\sum[(pi)\times\log(pi)] \quad \text{..... Eq. 1}$$

Genetic distance was estimated for the phylogenetic tree using the "C.S. Chord 1967" distance (Cavalli-Sforza and Edwards, 1967), followed by phylogeny reconstruction using Neighbor-Joining (NJ) as implemented in Power Marker and TreeView. The NJ tree was bootstrapped 1,000 times. A model-based Structure software (Falush et al. 2003; Pritchard et al. 2000) was used to infer population structure using a burn-in duration of 20,000, a run length of 20,000, and a model allowing for admixture and correlated allele frequencies, with testing for K = 2 to K = 10. Five independent runs produced consistent findings. To calculate the exact value of ΔK, the results were put into the Structure Selector software. (https://lmme.qdio.ac.cn/StructureSelector/)

Vegan package of R software was employed for Principal coordinate analysis (PCoA). Using the presence–absence data of forty-two SRAP primer combinations, Jaccard's dissimilarity coefficient was derived. Silhouette method was inferred to identify the optimal number of clusters, which was then used to determine Clusters using k-means clustering on the first two PCoA axes. Cluster boundaries were visualized using convex hulls.

RESULTS

A total of 7 Viola species were collected from the diverse KP regions of Pakistan, shown in Table 1. All species were identified based on morphological characteristics and submitted to the Herbarium. Preliminary observations indicated visible variation in morphology and habitat types, suggesting underlying genetic diversity, which was further assessed through molecular analysis. Seventeen Viola genotypes produced 456 polymorphic bands in total, indicating 100% polymorphism.

Statistics of SRAP primer amplification

The number of polymorphic bands assessed per primer combination ranged from seven to sixteen, with an average of 10.86 bands. The highest (16) bands were produced by the primer combination Me3Em4, followed by Me2Em5 and Me5Em3, 14 bands each. Me6Em3, Me3Em7, Me1Em4, and Me4Em5 all contained 13 polymorphic bands, while Me5Em2 and Me2Em7 had just seven (Table 3). The highest PIC value (0.32) was obtained for the Me2Em7 and Me3xEm6 primer pairs, followed by 0.31 for Me2Em3, Me4Em7, and Me6Em7, while the mean PIC value for all the markers was 0.25. The primer pair Me4Em4 had the lowest PIC value of 0.17. The major allele frequency ranged from 0.68 to 0.86 with an average of 0.79, and the minor allele frequency was between 0.14 and 0.32 with an average of 0.21 (Fig. 1). The average gene diversity was 0.30, varying from 0.21 (Me4Em4) to 0.41 (Me2Em7 and Me3Em6). The average Shannon's index was 1.88, ranging from 0.03 to 2.81. The highest Shannon's index value (2.81) was obtained for the Me1Em4 primer pair (Table 3).

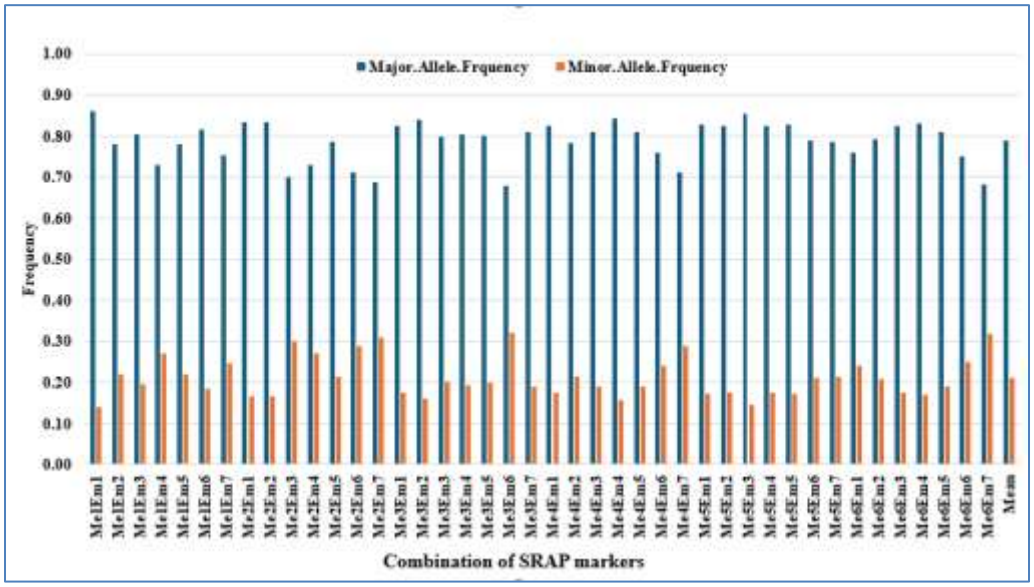

Fig. 1: Allele Frequency of the combination of 42 SRAP markers

Table 3: Summary statistics revealed by different primer pairs used in SRAP analysis

Primer pair	Total bands No. (TNB)	Polymorphic bands No.	Gene diversity	Polymorphism information content (PIC)	Shannon's index
Me1Em1	12	12	0.22	0.18	1.70
Me1Em2	8	8	0.30	0.25	1.44

Me1Em3	9	9	0.30	0.25	1.53
Me1Em4	13	13	0.35	0.28	2.81
Me1Em5	11	11	0.31	0.25	2.00
Me1Em6	8	8	0.27	0.23	1.26
Me1Em7	10	10	0.35	0.28	2.05
Me2Em1	12	12	0.26	0.22	1.77
Me2Em2	11	11	0.25	0.21	1.59
Me2Em3	10	10	0.39	0.31	2.38
Me2Em4	12	12	0.36	0.29	2.62
Me2Em5	14	14	0.30	0.25	2.03
Me2Em6	10	10	0.38	0.30	2.32
Me2Em7	7	7	0.41	0.32	1.74
Me3Em1	12	12	0.27	0.23	1.84
Me3Em2	10	10	0.24	0.21	1.46
Me3Em3	10	10	0.30	0.25	1.56
Me3Em4	16	16	0.30	0.25	2.70
Me3Em5	12	12	0.27	0.22	2.11
Me3Em6	9	9	0.41	0.32	2.27
Me3Em7	13	13	0.30	0.25	2.18
Me4Em1	11	11	0.25	0.21	1.61
Me4Em2	10	10	0.28	0.23	2.05
Me4Em3	9	9	0.27	0.22	1.41
Me4Em4	9	9	0.21	0.17	2.05
Me4Em5	13	13	0.26	0.21	2.02
Me4Em6	12	12	0.34	0.28	2.21
Me4Em7	10	10	0.38	0.31	2.33
Me5Em1	12	12	0.27	0.23	1.82
Me5Em2	7	7	0.26	0.22	1.05
Me5Em3	14	14	0.24	0.21	1.61
Me5Em4	10	10	0.26	0.21	1.64
Me5Em5	11	11	0.26	0.22	1.94
Me5Em6	12	12	0.32	0.26	2.19
Me5Em7	11	11	0.32	0.26	2.01
Me6Em1	11	11	0.32	0.26	2.31
Me6Em2	11	11	0.30	0.24	1.92
Me6Em3	13	13	0.28	0.23	1.71
Me6Em4	8	8	0.26	0.22	1.19
Me6Em5	12	12	0.27	0.22	2.06
Me6Em6	11	11	0.34	0.27	2.44
Me6Em7	10	10	0.40	0.31	0.03
Mean	**10.86**	**10.86**	**0.30**	**0.25**	**1.88**
Total	**456**	**456**			

Phylogenetic Tree Analysis

A dendrogram was generated using the Neighbor-joining tree method and genetic distance (Supplementary Table 1), which divides the tree into five clusters (Fig. 2). Hierarchical cluster analysis effectively distinguishes Viola genotypes. Cluster-1 was an outlier, containing *Viola hirta* collected from Mushkpuri and exhibiting the least similarity among all clusters, due to its distinct genetic profile compared to other genotypes. Cluster 2 included *Viola pilosa* from Baragali, *V. odorata*, and *V. cucullata* from Shogran. The sub-clade encompassing *V. pilosa* from Baragali and *V. cucullata* from Shogran exhibited greater genetic similarities than *V. odorata*. The largest cluster-3 included *V. mirabilis* from Shogran, *V. pilosa* from Miandam, *V. biflora* from Shogran, and *V. odorata* from Thandiani, Nathiagali, Bahrain, and Abbottabad. A very close genetic similarity was observed among these genotypes.

This cluster was divided into three sub-clades: subclade-1 contains *V. pilosa* from Miandam and *V. odorata* from Thandiani; subclade-2 contains *V. odorata* from Nathiagali and *V. biflora* from Shogran; and subclade-3 contains *V. odorata* from Bahrain and Abbottabad. *V. mirabilis* from Shogran exhibited relatively lower genetic similarity in this cluster than other sub-clades. Subclades 2 and 3 shared more genetic similarities than subclade 1.

V. canescens from Abbottabad, *V. pilosa* from Ayubia, and Namlimera made Cluster 4. The subclade containing *V. pilosa*

from Namlimera and *V. canescens* from Abbottabad was more similar to each other than *V. pilosa* from Ayubia. Cluster 5 contained *V. biflora* from Swat, *V. pilosa* from Thandiani, and Chattar Plain. In this cluster, the sub-clade containing *V. pilosa* from Chattar Plain and *V. biflora* from Swat showed greater genetic similarity with each other than *V. pilosa* from Thandiani. Based on molecular data, the Viola genotypes were classified into 6 groups with a diversity of 0.39.

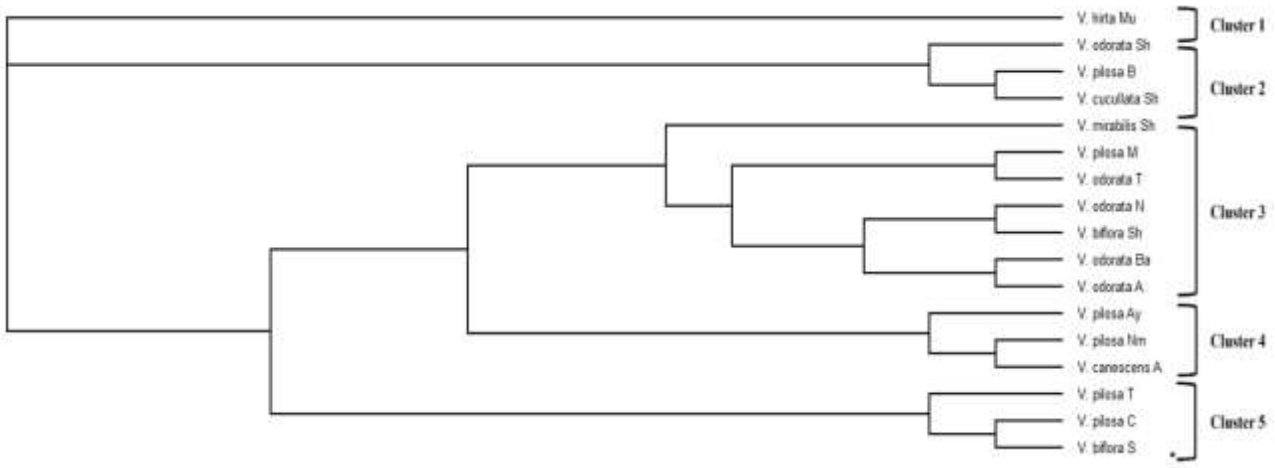

Fig. 2 Neighbor Joining Hierarchical Clustering Dendrogram Tree showing the genetic relationship between the 17 *Viola* genotypes.

Bayesian genetic structure

The structure analysis was performed for K values between 1 and 10. According to the Evanno method, the highest peak of Delta K was at K=2, suggesting the presence of two genetic populations (Fig. 3 a). The bar plot for K=2 showed that the Viola genotypes were divided into two clusters, colored red and green. Six Viola genotypes (*V. pilosa*_B, *V. cucullata*_Sh, *V. pilosa*_T, *V. odorata*_Sh, *V. hirta*_Mu, and *V. biflora*_S) have high membership probability to Cluster I (red), while the seven genotypes (*V. odorata*_A, *V. pilosa*_M, *V. odorata*_N, *V. odorata*_T, and *V. odorata*_Ba) were predominantly Cluster II (green) (Fig. 3 b). Some genotypes were partly assigned to these clusters, suggesting the presence of genetic admixture. Out of seventeen-four genotypes (*V. canescens*_A, *V. pilosa*_Nm, *V. pilosa*_C and *V. mirabilis*_Sh) were observed as having almost equal contribution of both clusters, indicating some degree of shared ancestry between the study species.

Principal Coordinates Analysis

Principal Coordinates Analysis (PCoA) was used to visualize genetic relatedness among different samples. The first two PCoA axes gave 38.5% of the total variation. Three distinct clusters were identified. Interestingly, cluster C3 included mainly *V. odorata* species collected from different sites, indicating the high level of genetic similarity among these samples. Cluster C2 comprised of the *V. pilosa* species. Cluster C1 showed a relatively divergent pattern as it comprised samples of diverse species, though *V. pilosa* were in the majority (Fig. 4).

DISCUSSION

The current study aimed to identify and report the genetic diversity of Viola species present in the northern areas of Pakistan. For this, 17 samples comprising different species of Viola collected from different sites were identified, collected, and evaluated at the molecular level by employing 42 distinct SRAP primer combinations. SRAP marker system revealed higher polymorphism, PIC values in different Viola genotypes, and hence proved to be a powerful tool and more informative in finding the genetic diversity (Li & Quiros, 2001; Li et al., 2009; Shao et al., 2010) that is useful for breeding. Literature showed that SRAP technology provides more information compared to other molecular markers, like AFLP, RAPD, ISSR, and SSR (Budak et al., 2004; Li wang et al., 2008; Ferriol et al., 2003; Kumar and Agrawal, 2017).

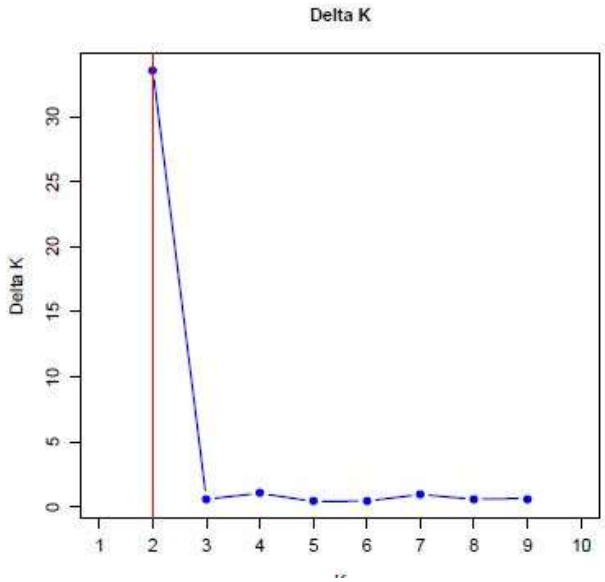

Fig. 3a: Graph showing the likelihood of Delta K=2

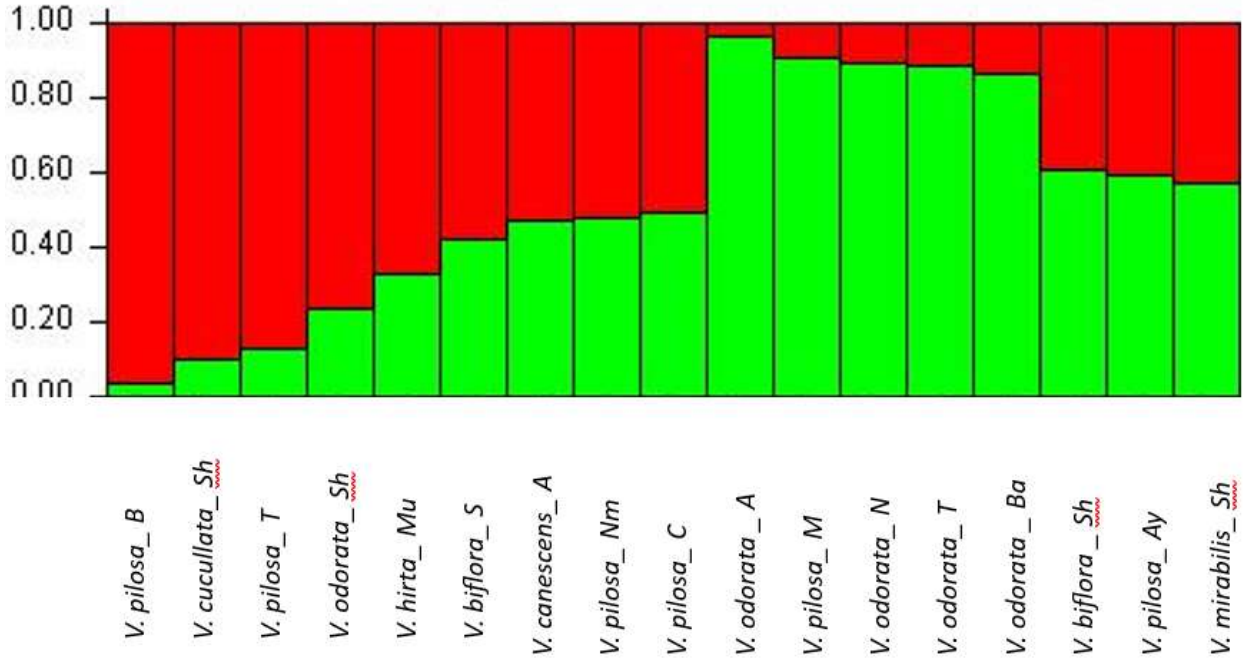

Fig. 3b: The Bar plot of Population structure analysis for 17 *Viola* genotypes at K=2

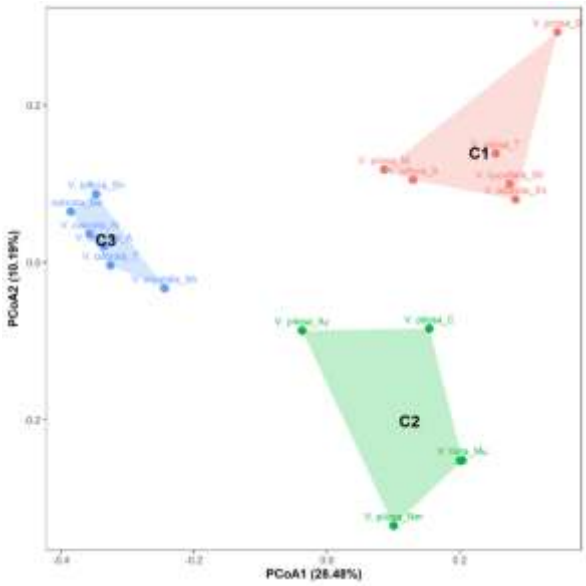

Fig. 4: Principal Coordinates Analysis showing the genetic relatedness between the 17 *Viola* genotypes in two components

Forty-two SRAP primers amplified 456 bands in total, among the 17 Viola genotypes, of which all bands were polymorphic with a mean percentage polymorphism (100%). This indicates that all the primer combinations used in this study were polymorphic and produced valuable results. However, this value was found to be higher than the value reported by Bhatt et al., (2017) in cumin (SRAP: 92%) and Wu et al., (2014), in Dioscorea species (SRAP: 93.5%), Chen et al., (2014), in *Mangnolia wufengensis* (SRAP: 88.0 %), Liu et al., (2013), (SRAP: 88.6%), Golkar and Nourbakhsh, (2019), in *Nigella sativa*, (SRAP: 83%), Zhou et al., (2021), in oil palm (*Elaeis guineesis* jacq.) (SRAP: 94.75 %). Kumar and Agrawal, 2017, in *Simarouba glauca* (SRAP: 30%). The PIC value of the SRAP markers ranges from 0.17 to 0.32 with an average of 0.25. The PIC values reported in different studies of various crops showed highly fluctuating values so the PIC value observed in our study was lower than those reported in different other crops like *Viola canescens*, *Nigella sativa*, (Kumar and Sharma, 2025; Golkar and Nourbakhsh, 2019), while higher than that reported in *Simarouba glauca*, oil palm, and *Centipede grass* (Kumar and Agrawal, 2017; Wang et al., 2023). A study demonstrated the utilization of 20 primer combinations to assess genetic diversity in hull-less barley varieties. A study by Yu et al. (2011) employed 11 SRAP primer combinations to evaluate genetic diversity in broccoli. The relationship among different genotypes was elucidated through Neighbor-Joining-based clustering patterns. Based on molecular data, the Viola genotypes were classified into 6 groups with a diversity of 0.39. The result is, according to Kermani et al. (2009), that a total of 4 clusters were generated in the cumin plant. Parashar & Malik (2014) reported 2 major groups in cumin by using different marker systems. A separate study indicated that two groups utilizing SRAP markers in Passion fruit exhibited a diversity of 0.093 (Oluoch et al., 2018). Cluster Analysis, according to molecular data, revealed slight level of species dependence in cluster formation like 4 out of 7 (57%) genotypes in cluster 3 comprised of *V. odorata* species, 2 out of 3 (67%) genotypes in cluster 4 and 5 comprised of *V. pilosa* species indicating a certain level of genetic similarity among different accessions of the same species. This trend is in accordance with previous studies. The variation among different accessions within and among clusters could be due to the environmental variables that also have a significant impact (Vilatersana et al., 2005) or genetic variations among different species like safflower and wheat (Najaphy et al., 2012). The involvement of a common fraction, convergent evolution, and subsequent selection may be the cause of similarities in genotypes classified in a single group by molecular cluster analysis (Reeves et al., 2012). According to cluster analysis, it was also revealed that there is a great variation in *Viola hirta* collected from Mushkpuri as compared to other genotypes, as it is an outlier that is in cluster-1. The genetic distance value of *Viola hirta* is also higher than that of other species, which shows higher genetic variation than other genotypes of Viola. The observed differences throughout the study could be attributed to the distribution of

genetic diversity, the type and number of molecular markers used, ecological forces acting among distinct populations that lead to the accumulation of genetic variations, and eco-geographic factors (El Harfi et al., 2021). Plants from Mushkpuri differ genetically from plants in any other place for a variety of reasons, including height, climate circumstances, physical differences in the soil, and distance from the source (Adhikari et al., 2012; Govinda Raj et al., 2015).

The Population Structure Analysis was performed for 17 Viola genotypes based on SRAP markers data, revealing a maximum likelihood of ΔK =2. Various other investigations in different plants revealed the same value of ΔK, like a study on 83 Cabbage landraces in China, using AFLP markers data (Kang et al., 2011), in Ziziphus species by using SSR markers (Sareen et al., 2023), in Indian blackberry (*Syzygium cumini* L.) by using CBDP and SCoT markers (Kumar et al., 2025). Interestingly, the structure analysis also showed a certain level of similarity among different genotypes of the same species, as *V. odorata* collected from different sites except Shogran. These results indicate that geographic and environmental variations have caused genetic diversity within the same species; intra-species genetic conservation was also observed to a certain level. PCoA analysis revealing three clusters with a certain level of species-level conservation indicated that variation within species is minimal, showing that genetic variations are not strongly site/ecosystem dependent. This study provides useful insight about pattern of genetic diversity among and within different Viola species present in the northern areas of Pakistan.

CONCLUSION

Our results conclusively revealed the power, effectiveness, and informativeness of SRAP markers for finding genetic diversity; they were efficient tools. High detection of polymorphism among collected Viola species is offered by the SRAP marker from Pakistan's northern regions. Although geographic and environmental variations have caused genetic diversity within the same species, intra-species genetic conservation was also observed to a certain level. Pharmaceutical applications, food, and valuable genetic resources for future breeding are highlighted by divergence patterns and distinct clustering, particularly in *Viola hirta*. In the herbal sectors and in the pharmaceutical sectors, these findings can be used as raw material to increase the productivity and income of farmers.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest

AUTHORS' CONTRIBUTION STATEMENT

SMS and RA conceived the idea, supervised research, and wrote the manuscript. SL performed the experiment, curated the data, and wrote the manuscript. AN and ARK assisted in reviewing and revising the final draft. ARK and SMS performed data analysis and wrote the discussion chapter. All authors approved the final draft of the manuscript.

ACKNOWLEDGEMENTS

The authors acknowledge the Higher Education Commission of Pakistan for providing the funds under NRPU Research Project No.9857. This article is part of the PhD dissertation of the first author.

REFERENCES

Abolghasemi, S., Naderi, R., & Fattahi, M. M. R. (2020). Evaluation of genetic diversity in Iranian Violet (Viola spp) populations using morphological and RAPD molecular markers. J Genet Resour 6(2): 157-171.

Adhikari D, Kaneto M, Itoh K, Suyama K, Pokharel BB, Gaihre YK (2012). Genetic diversity of soybean-nodulating rhizobia in Nepal in relation to climate and soil properties. Plant Soil. 357: 131–145

Aneja B, Yadav NR, Chawla V, Yadav RC (2012). Sequence-related amplified polymorphism (SRAP) molecular marker system and its applications in crop improvement. Mole. Breed 30: 1635-1648

Batiha GE, Lukman HY, Shaheen HM, Wasef L, Hafiz AA, Conte-Junior CA, Al-Farga A, Chamba MVM, Lawal B (2023). A Systematic Review of Phytochemistry, Nutritional Composition, and Pharmacologic Application of Species of

the Genus Viola in Noncommunicable Diseases (NCDs). Evid Based Complement Alternat Med. 5406039. doi: 10.1155/2023/5406039. PMID: 37941895; PMCID: PMC10630019.

Bhatt J, Kumar S, Patel S, Solanki R (2017). Sequence-related amplified polymorphism (SRAP) markers based on genetic diversity analysis of cumin genotypes. Ann. Agrar. Sci.15(4): 434-438

Cavalli-Sforza, LL, Edwards, AW (1967). Phylogenetic Analysis. Models and Estimation Procedures. The American Journal of Human Genetics,19, 233-257.

Dastagir G, Samra B, Noor Ul Uza, Rainer W. Bussmann, Imran A, Samiullah (2023). Microscopic evaluation, ethnobotanical and phytochemical profiling of a traditional drug *Viola odorata* L. from Pakistan. Ethnobotany Research and Applications,25, 1–24.

Doyle J (1991) DNA protocols for plants. In Molecular techniques in taxonomy (pp. 283-293). Berlin, Heidelberg: Springer Berlin Heidelberg.

El Harfi M, Charafi J, Houmanat K, Hanine H, Nabloussi A (2021). Assessment of genetic diversity in Moroccan sesame (*Sesamum indicum*) using ISSR molecular markers. Oilseeds and fats, Crops and Lipids,28: 3.

El-Shazly HH, Ahmed HIS, Hamouda MM, Badr A (2020). Genetic diversity and population structure of the medicinal plant *Achillea fragrantissima* (Forssk.) Sch. Bip. in the mountains of South Sinai, Egypt. Plant Gene. 21: 100212.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 164(4):1567-87. doi: 10.1093/genetics/164.4.1567. PMID: 12930761; PMCID: PMC1462648.

Fatima G, Rida B, Rimsha D (2023). Antioxidant and Antibacterial Potential of *Achillea millefolium, Sisymbrium irio* and *Viola canescens* from Azad Jammu Kashmir, Pakistan: Potential of *Viola canescens* from Azad Jammu Kashmir. Biological Sciences - PJSIR, 66, 1-9.

Ferriol M, Pico B, Nuez F (2003). Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. Theor. Appl. Genet. 107: 271-282.

Flicker B, Ballard HE (2015). Afrohybanthus (Violaceae), a new genus for a distinctive and widely distributed old-world hybanthoid lineage. Phytotaxa. 230: 39–53.

Golkar P, Nourbakhsh V (2019). Analysis of genetic diversity and population structure in *Nigella sativa* L. using agronomic traits and molecular markers (SRAP and SCoT). Ind. Crop. Prod. 130: 170-178.

Hussain S, Murtaza G, Qureshi RA (2016). Floristic studies of angiosperms of Rawalakot Azad Jammu and Kashmir, Pakistan. JAPS.26(6).

Kang J, Fang Z, Wang X, Xu D, Liu Y, Yang L, Zhuang M, Zhang Y (2011). Genetic diversity and relationships among cabbage (*Brassica oleracea* var. capitata) landraces in China revealed by AFLP markers. Afr. J. Biotechnol.10(32): 5940-5949

Kermani M, Maraashi SH, and Safarnejad A (2009). Investigation of genetic variation within and among two species of Cuminum spp. using AFLP markers. Iran J Rangelands For Plant Breed. Genet. Res. 16(2): 198-206

Kumar A, Singh K, Singh AK, Prakash J, Goswami AK, et al. (2025). Genetic diversity and population structure analysis of Indian blackberry (*Syzygium cumini* L.) using CAAT box-derived polymorphism (CBDP) and start codon targeted polymorphism (SCoT) markers. J. Genet. Eng. and Biotechnol.23(1): 100468

Kumar J, Agrawal V (2017). Analysis of genetic diversity and population genetic structure in *Simarouba glauca* D C. (an important bio-energy crop) employing ISSR and SRAP markers. Ind. Crop. Prod.100: 198-207.

Kumar R, Sharma V (2025). Elucidation of genetic diversity and population structure in medicinal plant *Viola canescens* from Northwestern Himalayas. Journal of Applied Research on Medicinal and Aromatic Plants, 100657.

Li Y, Fan X, Shi T, Zhang Q, Zhang Z (2009). SRAP marker reveals genetic diversity in tartary buckwheat in China. Front. Agric. China.3: 383-387.

Li, G. Quiros CF (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. Theor. Appl. Genet.103: 455-461

Liu K, Muse SV. (2005). PowerMarker: Integrated Analysis Environment for Genetic Marker Data. Bioinformatics.21, 2128-2129. http://dx.doi.org/10.1093/bioinformatics/bti282

Liu L, Yang Z, Wei S, Ouyang Z, Wu K, Peng Z (2012). ISSR and SRAP markers in the genetic relationship analysis among Pinellia in China. J. Med. Plants. Res.6(19): 3596-3602.

Madhuri Parashar MP, Malik CP (2014). Appraisal of genetic diversity in *Cuminum cyminum* L. using molecular markers. Indian J.3(3): 143-156

Najaphy A, Parchin RA, Farshadfar E, Farshadfar E (2012). Comparison of phenotypic and molecular characterizations of some important wheat cultivars and advanced breeding lines. Aust. J. Crop Sci.6(2): 326-332

Nam VT, Hang PLB, Linh NN, Ly LH, et al. (2020). Molecular markers for analysis of plant genetic diversity. Vietnam J. Biotechnol.18(4): 589-608

Oluoch P, Nyaboga EN, Bargul JL (2018). Analysis of genetic diversity of passion fruit (*Passiflora edulis* Sims) species grown in Kenya by sequence-related amplified polymorphism (SRAP) markers. Ann. Agrar. Sci.16(4): 367-375.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genet.155(2): 945-959.

Reeves PA, Panella LW, Richards CM (2012). Retention of agronomically important variation in germplasm core collections: implications for allele mining. Theor. Appl. Genet.124: 1155-1171

Robarts DWH, Wolfe AD (2014). Sequence-related amplified polymorphism (SRAP) markers: A potential resource for studies in plant molecular biology. Appl. Plant Sci.2(7): 1400017.

Sareen A, Sharma V, Gupta RC (2023). Assessment of genetic diversity and population structure in wild Ziziphus species from northwest India using SSR marker technique. J. Genet. Eng. and Biotechnol.21(1):4

Sidhu G, Banks T, Wolyn D (2025). Genetic diversity and population structure analysis in *Asparagus officinalis*. J. Genet. Eng. and Biotechnol.23(2): 100491.

Vilatersana R, Garnatje T, Susanna A, Garcia-Jacas N (2005) Taxonomic problems in Carthamus (Asteraceae): RAPD markers and sectional classification. Bot. J. Linn. Soc.147(3): 375-383

Wahlert GA, Hoyos-Goméz SE, Ballard HE (2018). Systematic studies in *Neotropical Rinorea* (Violaceae): two new sections and a new generic segregate. Brittonia.70: 140–147

Wang S, Shen C, Zhang S, Di H, Wang Y, Guan F (2025). A Comprehensive Review of the Phytochemistry and

Therapeutic Efficacy of *Viola yedoensis* Makino. Molecules. 30(9):1922. https://doi.org/10.3390/molecules30091922

Wang X, Gou W, Wang T, Xiong Y, Xiong Y, Yu Q, Dong Z, Ma X, Liu N, Zhao J (2023) Genetic diversity analysis and molecular characteristics of wild centipede grass using sequence-related amplified polymorphism (SRAP) markers. Peer J 11: e15900

Zhou L, Yarra R, Cao H, Zhao Z (2021). Sequence-related amplified polymorphism (SRAP) markers based genetic diversity and population structure analysis of oil palm (*Elaeis guineensis* Jacq.). Trop. Plant Biol.14: 63-71.

Supplementary Table 1: Genetic Distances between 17 *Viola* genotypes

out	V. biflora_S	V. biflora_Sh	V. canescens_A	V. cucullata_Sh	V. hirta_Mu	V. mirabilis_Sh	V. odorata_A	V. odorata_Ba	V. odorata_N	V. odorata_Sh	V. odorata_T	V. *pilosa*_Ay	V. *pilosa*_B	V. *pilosa*_C	V. *pilosa*_M	V. *pilosa*_Nm	V. *pilosa*_T
V. biflora_S	0.0000																
V. biflora_Sh	0.2732	0.0000															
V. canescens_A	0.2963	0.2886	0.0000														
V. cucullata_Sh	0.2366	0.2289	0.2174	0.0000													
V. hirta_Mu	0.2789	0.2789	0.2636	0.2078	0.0000												
V. mirabilis_Sh	0.3040	0.2655	0.3078	0.2212	0.2328	0.0000											
V. odorata_A	0.3482	0.3174	0.3444	0.3347	0.3540	0.3251	0.0000										
V. odorata_Ba	0.3097	0.2520	0.3367	0.2924	0.3116	0.2943	0.2539	0.0000									
V. odorata_N	0.3251	0.2482	0.3367	0.2655	0.3193	0.3097	0.3116	0.2501	0.0000								
V. odorata_Sh	0.2289	0.2597	0.2559	0.1885	0.2078	0.2405	0.3463	0.2963	0.3193	0.0000							
V. odorata_T	0.3059	0.2905	0.3367	0.3116	0.3193	0.3444	0.3386	0.3386	0.3078	0.3424	0.0000						
V. pilosa_Ay	0.2501	0.2924	0.2616	0.2559	0.2789	0.2809	0.3597	0.3482	0.3213	0.2482	0.2905	0.0000					
V. pilosa_B	0.2020	0.2135	0.2212	0.1347	0.1962	0.2212	0.3232	0.2539	0.2655	0.1616	0.2963	0.2443	0.0000				
V. pilosa_C	0.2347	0.3232	0.2655	0.2482	0.2674	0.3078	0.3905	0.3444	0.3213	0.2789	0.3444	0.2732	0.2366	0.0000			
V. pilosa_M	0.3386	0.3540	0.3732	0.3146	0.3328	0.3232	0.3762	0.3867	0.3674	0.3328	0.3405	0.3694	0.3097	0.3366	0.0000		
V. pilosa_Nm	0.2712	0.2789	0.2251	0.2424	0.2193	0.2789	0.3655	0.3270	0.3040	0.2655	0.3116	0.2597	0.2520	0.2866	0.3405	0.0000	
V. pilosa_T	0.2078	0.2424	0.2462	0.1751	0.1981	0.2693	0.3559	0.2828	0.3174	0.2174	0.2712	0.2616	0.2481	0.2116	0.3040	0.2174	0.0000