

Original Research Article

Genetic Diversity Analysis and Molecular Diversity in Greengram (*Vigna radiata* (L.) Wilczek) Genotypes using SSR Markers

Need to revise the title in accordance with the manuscript “Have both field and Marker data”

ABSTRACT

A study on the genetic and molecular diversity of 40 genotypes of Greengram was conducted using SSR markers. The experimental material contained significant genetic diversity; nevertheless, for all yield-related and yield-attributing features, phenotypic coefficient of variation exceeded genotypic coefficient of variation. The genotypes LGG 574 (8.80), PDM 139 (8.34), and Pant Mung 6 had the highest seed yields (7.78). The highest observed cluster distances between clusters 5 and 6 (472.88) and clusters 4 and 5 (432.89). Among all the factors PC 1 to PC 10, the PC 1 (19.99) accounted maximum proportion of variability in the set of all variables and remaining components accounted for progressively lesser and lesser amount of variation. The first six principal components (PC-1 to PC-6) with eigen values of 2.94, 2.62, 2.18, 1.40, 1.28, and 0.86, respectively, accounted for 86.95% of the total variance for all the qualities, according to principal component analysis. The genotypes 23, 32, 33, and 21 were spread out relatively far from other genotypes in the scatter plot, suggesting that they might be different from other genotypes. PIC values of 10 SSR loci, where the VR 86 marker produced the greatest PIC Value percentage. In comparison to Cluster II, III, IV, and V, Cluster I has the most genotypes (25), and the use of SSR markers in this work to differentiate between genotypes was made possible by the high polymorphism information richness of this cluster.

Rewrite including SSR Findings

Keywords: Greengram, GCV, PCV, SSR, PCA, genetic diversity, molecular diversity.

INTRODUCTION

Mungbean, often known as Greengram, is a member of the Leguminosae family (*Vigna radiata* (L.) Wilczek, $2n=22$). Crop pollination in Greengram happens naturally, although cross pollination is also present, albeit in very small amounts. India is the world's major exporter and user of Greengram. There is a lot of diversity among Greengram cultivars and within the Greengram species because it is a self-pollinated crop (**Bisht *et al.*, 2005**). Greengram has a special role in the diversification of Indian agriculture and the fight against malnutrition among the nation's vegetarian population. Greengram is a short-lived legume crop produced for its high protein (25%) and lysine (504 mg/g) content. It is thought to be a cure for Beriberi and is high in vitamin B. The area of this crop in India is 5.20 million hectares, and its productivity is 2.97 million tonnes and 572 kg per ha. (Unknown **2020–2021**). The biggest producers of Greengram in India are the states of Rajasthan (19.69 lakh ha; 48.65 lakh acres), Karnataka (4.14 lakh ha; 10.23 lakh acres); Maharashtra (3.68 lakh ha; 9.10 lakh acres); Madhya Pradesh (1.60 lakh ha; 3.95 lakh acres); and Telangana (0.55 lakh ha; 1.35 lakh acres) (**pjtsau.edu.in 2021-22** forecast of Greengram).

Before beginning an effective breeding programme, it is essential to conduct research on genetic variation, including heritability, phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), and genetic progress. Because there is greater variation in segregating generations that can be used to develop a cultivar, genetic divergence between the parents is necessary (**Nimbalkar *et al.*, 2017**). Examining population variability and developing our understanding of it creates the foundation for efficient and productive breeding operations (**Bello *et al.*, 2012**). Heritability estimates alone are frequently less informative than genetic progress and heritability estimates for predicting gain under selection (**Johnson *et al.*, 1955**). Breeders must study genetic variation in genetic resources in order to develop ways for incorporating valuable diversity into their breeding programmes, better comprehend the evolutionary and genetic linkages among populations, and select germplasm in a more methodical and effective manner (**Lavanya *et al.*, 2008**). The most crucial stage of any crop development endeavour continues to be the evaluation of the genetic variety that currently exists in the basic gene pool utilizing multivariate analytic techniques like D^2 statistics and factorial analysis. The relative significance and utility of various variables and genotypes in a data set are assessed using principal component analysis. It reduced a huge number of associated variables to a small number of independent main components, which accounted for the majority of the data variance. Molecular markers make it easier to spot labelling mistakes, track down the rightful owner of a disputed cultivar, and regularly recognize cultivars in

nurseries. Use of the molecular marker simple sequence repeat (SSR) has been made to evaluate the genetic diversity in Greengram (**Gwag *et al.*,2010**). The study of molecular markers is essential to genomic research. Because of their repeatability, multiallelic nature, codominant inheritance, relative abundance, and good genomic coverage, SSRs stand out among other marker systems like restriction fragment length polymorphism (RFLP), RAPD, sequence tagged sites (STSs), and AFLP. SSRs are collections of brief tandem repeats of nucleotide bases found all over the genome. The abundance of SSRs in the genomes that have been studied so far and their hypervariability are key characteristics that contribute to their popularity (**Singh *et al.*, 2013**).Therefore, any breeding effort can benefit from the genetic diversity and divergence found in the materials. In order to improve the genotypes, we would further benefit from the assessment of variance because it would provide us an accurate picture of the level of variation. In order to assess the genetic diversity and linkages among genotype lines, the primary goal of this work was to describe Greengram genotypes using molecular markers.**ADD current status of SSR marker used in this crop**

MATERIALSANDMETHODS

A study on the genetic and molecular diversity of 40 genotypes of Greengram was conducted using SSR markers at the Department of Plant Breeding and Genetics, SHUATS, Prayagraj, during the kharif 2021 in a randomized block design followed by molecular work at Division of Plant biotechnology, Indian Institute ofPulses Research (IIPR), Kanpur. The recorded morphological data were subjected to statistical studies of genetic and genomic diversity. To collect data on plant height, number of primary branches, number of clusters, number of pods, pod length in cm, number of seeds pod, number of seeds per pod, number of seeds per pod, number of seeds per pod, 100 seed weight per pod, biological yield, harvest index, and seed yield plant in Greengram, five randomly chosen plants were chosen. The agronomic trait observations would aid in assessing genetic diversity based on phenotypic data, allowing one to compare genetic diversity as explained by phenotypic data with genotypic data. The distance between two populations, D^2 , as defined by Mahalanobis (1936), was determined using Tocher's approach, which Rao outlined (1952). The Singh and Chaudhary approach was used to evaluate the contribution of individual characters to divergence (1985). PCA was performed usingXLstatver 2021 software to assess genotype divergence in terms of spatial distance in a two-way graphical graph setting the relative location of each genotype and to determine the contribution of characters to the overall variability for all the characteristics under consideration.

Genomic DNA extraction and quantification

Total genomic DNA was extracted from 40 genotypes using the Cetyl Trimethyl Ammonium Bromide (CTAB) extraction technique (Doyle and Doyle, 1987) and spectrophotometrically measured using a nano spectrophotometer (Implen, Germany). The concentration of genomic DNA of each genotype was measured by comparing it to a known concentration of DNA of 200ng and adjusting it to 20ng μl^{-1} by diluting with Milli-Q water for SSR marker analysis.

Accumulation of the SSR-PCR Screening of the Greengram germplasm lines shown in Table 2 was conducted using thirty SSR or microsatellite repeat primers. PCR amplification was performed in a 10 μl PCR reaction contains 0.3 μl of dNTPs, 0.2 μl of 3 units *Taq* DNA polymerase, 1 μl of 10x *Taq* buffer, 2 μl of DNA, 1 μl of primer mix and 5.5 μl of Milli-Qwater. The amplification was carried out under reaction conditions of pre-denaturation at 94 °C for 5 minutes, followed by 40 cycles of denaturation at 94 °C for 40 seconds, annealing at 48- 61.8 °C for 1 minute, extension at 72 °C for 60 seconds, and final extension for 10 minutes at 72 °C with a hold temperature of 4 °C. Electrophoresis was utilized to separate amplification products using 3% metaphor agarose (Sigma-Aldrich, India), and gel pictures were taken.

Data analysis

Amplicons generated using SSR primers and resolved on agarose gels for each marker allele-genotype combination were scored quantitatively for presence in a binary coding, i.e., presence was marked as (1) and absence as (0). Using the Gene Alex 6.5 programme, the binary data was analysed for the number of alleles, Shannon's information index, PIC (Polymorphism information content), and heterozygosity. The data for PCA (PrincipalComponent Analysis) was collected using the XLstatver 2021 application. For 40 Greengramgenotypes, the NTsysprogramme was used to produce a molecular dendrogram cluster analysis using the ward technique and squared Euclidean distance.

RESULTS AND DISCUSSION:

In the current study, analysis of variance revealed a significant genetic difference for each trait. This implies that there was enough room in the present gene pool to choose promising lines for yield and component characteristics. The appearance of high variability may be due to the utilization of several data sources as well as environmental factors that modify the phenotypes. In terms of biological yield, the genotypes LGG 574 (8.80) had the maximum seed yield,

followed by PDM 139 (8.34), Pant Mung 6 (7.78), and IPM 99-125 (7.40). LGG 574 has the most primary branches and clusters, according to PDM-11. The estimated genotypic variance and phenotypic variance showed that, for all yield-related and yield-attributing characters, phenotypic variation was higher than genotypic variance, indicating the influence of environmental variables on these traits. It was discovered that the phenotypic coefficient of variation was greater than the genotypic coefficient of variation, which raises the possibility that the environment may affect how the feature under study manifests. **Ramakrishnan et al.** reported a similar discovery (2018). The production of seeds, the number of clusters per plant, and the number of pods per plant showed the most notable differences between GCV and PCV. **Nitesh et al.** observed a similar finding (2017). The highest heritability was found in the following categories: days to maturity (83.254), plant height (80.163), number of primary branches (88.552), number of pods per plant (67.602), number of seeds per plant (77.891), seed index (98.782), harvest index (77.847), biological yield per plant (66.808), and seed yield per plant (76.097), respectively. **Muthuswamy et al.** also reported similar results (2019). The highest estimates of genetic advance as a percentage of mean were found for plant height (26.224), number of primary branches (68.006), number of pods per plant (28.558), number of seeds per plant (21.061), number of seeds per pod (21.808), number of seeds per plant (seed index (32.93), biological yield per plant (23.633), and number of seeds per plant (34.474). The data acquired on 13 yield and yield contributing variables for the 39+1(check) genotypes of Greengram were subjected to genetic divergence using the **Mahalanobis D²** statistic. The size of the D² values showed that there was a lot of genetic variety produced by the item under consideration. The 39+1 (check) genotypes were randomly assigned into six clusters using the Tocher technique, with Cluster I having the most genotypes (26), Cluster II having the fewest genotypes (1), and Cluster III, IV, V, VI, and VII having the fewest genotypes (1). Genetic diversity and geographic diversity do not correlate, as shown by the genotype distribution, suggesting that other factors, such as genetic drift, environmental variation, natural and artificial selection, and breeding material exchange, are more likely to be the cause of diversity than geographic isolation. According to the values of the intercluster distances between the seven clusters, cluster V and VI (472.88) had the highest divergence, followed by cluster IV and V (432.89), cluster V and VII (424.78), cluster I and V (300.20), cluster III and V (276.84), cluster II and VI (218.80), cluster II and VII (188.63), cluster II and IV (181.96), cluster III and IV (180.58), cluster III and VII (180.07), and (173.77). Comparable outcomes have been reported by **Suhelet et al. (2015)**, **Manoj and Sachin (2015)**, **Sai Rekha et al. (2015)**. In order to increase the ability of the two parents to combine and produce greater yield traits, it is advised that two

parents be chosen for a hybridization programme who differ more from one another. The current study revealed that the number of seeds produced per plant (20.30%) had the biggest contribution to total divergence, followed by harvest index (15.15%), biological yield (9.92%), seed index (8.65%), number of primary branches per plant (8.52%), and number of clusters per plant (6.55). Similar finding was observed by **Katiya and Kumar (2015)**, **Aijaz et al. (2016)**, **Gaurav et al. (2017)**, **Sen et al. (2017)**, and **Sharma et al. (2018)**. Thirty SSR primers were utilised for the molecular analysis, of which ten were polymorphic and the other twenty were monomorphic. Ten SSR markers were used in PCA for 40 genotypes of Greengram based on the findings of the polymorphic bands. In the current investigation, the first six principal components (PC-1 to PC-6) with eigenvalues of 2.94, 2.62, 2.18, 1.40, 1.28, and 0.86, respectively, accounted for 86.95% of the total variance for all the features. Characters' contributions to the divergence would increase with increasing absolute value in the PC (**Singh et al., 2017**). The first principal component (PC-1) accounted for 22.63% of the variability, with biological yield (0.42), seed yield per plant (0.40), and plant height being the most important contributing variables (0.36). Similarly, significant loadings of days to maturity (0.37) and days to 50% flowering (0.29), which explained 20.17% of total variation, were detected in the second principal component (PC-2). Pod length (0.48), days to 50% flowering (0.40), and days to maturity (0.39), together accounted for 16.78% of total variance in the third main component (PC-3). The fourth principal component (PC-4) accounted for 10.8% of total variance and was distinguished by significant loadings for number of clusters per plant (0.50), pod length (0.47), and number of seed per pod (0.38). Days to 50% pod setting (0.70) and harvest index (0.27) contributed 9.88% of total variation to the fifth principal component (PC-5), indicating that these variables might be employed effectively for germplasm selection for yield increase in Greengram. Principal components with eigenvalues greater than one should be retained as potential contributors to diversity (**Kaiser, 1960**). The sixth principal component (PC-6) contributed less to divergence, with an eigen value less than 1.0 and a variability of 6.64%. As a result, variables such as plant cluster number, plant pod number, plant height, days to maturity, days to 50% flowering, and seed yield plant-1 contribute the most to divergence. This finding corresponded to the findings of **Jakhar and Kumar (2018)** for plant height and **Mahalingam et al. (2019)** for seed yield plant. The fifth PC contributed less to divergence, with an eigen value less than 1.0 and a variability of 9.8%. The results agreed with those of **Jakhar and Kumar (2018)** and **Mahalingam et al. (2019)**. The first three main components accounted for 59.59% of the total variance with eigenvalues greater than one. In order to create a 2D scatter diagram, the scores of PCA for the first three principal components for 40 genotypes were calculated and

plotted onto a graph, with PCA I in the X-axis and PCA II in the Y-axis (Table 5, Fig. 1). PCA was performed for 30 SSR markers for 40 Green gram genotypes. Among all the factors PC1 TO PC10, the PC1 (19.995) accounted for maximum proportion of variability in the set of all variables and remaining components accounted for progressively lesser and lesser amount of variation. The PC1 accounted for maximum variability i.e., 19.995% and PC2 (16.562%), PC3 (13.595%), PC4 (12.929%), PC5 (10.041%), PC6 (8.824%), PC7 (6.849%), PC8 (5.615%), PC9 (4.343%), PC10 (1.247%). The highest Eigen value for PC1 (2.00) and the least is PC10 with 0.125 (Table 6). The first principal component PC1 showed high positive loading for VR80 (0.891) followed by CEDG271 (0.693), CEDG048 (0.280), VR48 (0.048) and VR91 (0.026), where it has showed high negative loading for CEDG 21 (-0.580), CEDG 236 (-0.430), CEDG300 (-0.310), VR (-0.120) and CEDG115 (0.100). The second principal component PC2 showed high positive loading for CEDG300 (0.647) followed by CEDG048 (0.629), VR91 (0.371) and CEDG236 (0.093). whereas it showed high negative loading for VR48 (-0.520), CEDG115 (-0.520), VR86 (-0.340), CEDG21 (-0.170) and CEDG271 (-0.130). The third principal component PC3 showed high positive loading for VR 86 (0.708) followed by CEDG236 (0.441), VR48 (0.332), CEDG048 (0.188), CEDG300 (0.101) and VR80 (0.076) whereas it showed high negative loading for CEDG21 (-0.460), CEDG115 (-0.400), VR91 (0.320) and CEDG271 (-0.180). The fourth principal component PC4 showed high positive loading for CEDG236 (0.604) followed by CEDG271 (0.541), CEDG21 (0.311), VR80 (0.283), VR91 (0.273), VR86 (0.199) and CEDG300 (0.142). whereas it showed high negative loading for CEDG048 (-0.460), VR 48 (-0.290), CEDG115 (-0.170). The fifth principal component PC5 showed high positive loading for CEDG 115 (0.536), VR91 (0.406), VR86 (0.256), CEDG236 (0.190) and CEDG048 (0.057), whereas it showed high negative loading for VR48 (-0.420), CEDG21 (-0.410), CEDG300 (-0.240), CEDG271 (-0.220) and VR80 (-0.010). The sixth principal component PC6 showed high positive loading for VR91 (0.712), VR48 (0.501), CEDG271 (0.075), VR86 (0.067), CEDG048 (0.053) and CEDG21 (0.015), whereas it showed high negative loading for VR80 (-0.210), CEDG115 (-0.200), CEDG300 (-0.160) and CEDG236 (-0.060). The seventh principal component PC7 showed high positive loading for CEDG300 (0.480), VR86 (0.369), CEDG115 (0.368), CEDG271 (0.208), CEDG048 (0.195), CEDG21 (0.172), VR48 (0.109), VR91 (0.046) and VR80 (0.026), whereas it showed high

negative loading for CEDG236 (-0.240). The eighth principal component PC8 showed high positive loading for CEDG 048 (0.398), CEDG21 (0.323), VR86 (0.226) and CEDG271 (0.149), whereas it showed high negative loading for CEDG300 (-0.370), VR48 (-0.220), VR80 (-0.140), CEDG115 (-0.110), VR 91 (-0.08) and CEDG236 (-0.020). The ninth principal component PC9 showed high positive loading for CEDG 236 (0.390), CEDG048 (0.272), CEDG115 (0.250), VR48 (0.224), CEDG271 (0.137) and VR21 (0.010), whereas it showed high negative loading for VR86 (-0.270), VR91 (-0.080), VR80 (0.020) and CEDG115 (-0.010). The tenth principal component PC10 showed high positive loading for VR 80 (0.237), CEDG21 (0.153), VR91 (0.051), VR48 (0.046), CEDG048 (0.046), CEDG236 (0.023), CEDG115 (0.012) and VR86 (0.007), whereas it showed high negative loading for CEDG300 (-0.020) and CEDG271 (-0.190). The genotypes 23, 32, 33, and 21 were located somewhat far from other genotypes, according to the 2D structure. Better recombinants would result from combining these genotypes (**Thippani et al., 2017**). The 40 Greengram genotypes are clustered using the ward technique and squared Euclidean distance utilising the 10 SSR markers. According to the dissimilarity coefficient, the 40 genotypes are separated into five clusters: Clusters I, II, III, IV, and V. Cluster I, which comprises 25 genotypes, is further subdivided into two more compact clusters. There are 4 genotypes in subcluster IB, compared to 21 in subcluster IA. The SSR marker CEDG 115 & VR 91 has the most alleles (5 total), followed by CEDG 300, VR 86, VR 48, VR 80, CEDG 236 and CEDG 048. The two alleles shared by CEDG 271 and CEDG 21 are the fewest, and the two alleles shared by CEDG 115 and VR 91 are the most (5 total). The isolation and fixation of particular alleles during reproduction may be the cause of high diversity (**Hartl and Clark, 1997**). PIC is a significant measure that determines a marker's capacity to distinguish between various genotypes and evaluates the efficacy of polymorphic loci. With values ranging from 0.049 to 0.694, the average polymorphic information content value (PIC) per locus was 0.427. Singh *et al.* report a similar finding (2013). The high (I) diversity result for VR 86 and the lowest (I) diversity result for CEDG 21 in the current study **Wang et al., (2018)** found a similar discovery, and the average Shannon's index for the 10 SSR polymorphism markers is 0.432. The random measurement of species variety is done using Shannon's information index (I). The fixation index has a mean of 0.774 and a range of 1.000 to -0.076. The high fixation index at most loci could be attributed to self-pollination in Greengram-positive genotypes with a small number of heterozygous loci. A high fixation index was discovered to be connected with low variance. **Changyouet et al., 2017**;

Wang *et al.*, 2017; Kanimoli *et al.*, 2018; Muthusamy *et al.*, 2008; Results from this study could be used to further improve crops using advance marker systems and would be very helpful in Greengram breeding projects.

UNDER PEER REVIEW

Table 1: Genotypes of Greengram with their pedigree and salient features

S. No.	Genotypes	Centre responsible for developing	Pedigree	Year of release	Salient features
1	PDM-11	IIPR, KANPUR	Selection from LM595	1987	Erect bushy, shining green seed, suitable for spring season
2	PDM-54	IIPR, KANPUR	Selection from Kundawa Bahraich local (UP)	1987	Erect bushy, shining green seed
3	ADT-3	TNAU, ADUTHARAI	(M70-16 × Rajendra) × G65	1991	Semi spreading, resistant to MYMV
4.	PUSA 9072	IARI, NEW DELHI	PUSA 106 × 10-215	1995	Moderately resistant to Powdery Mildew
5	Pant Moong 4 (UPM92-1)	GBPAU, PANTNAGAR	T-44 × UPU-2	1997	Erect, Dull green seed, resistant to MYMV
6	PUSA 9531	IARI, NEW DELHI	Selection from NM9473	2000	Resistant to MYMV
7	PUSA VISHAL	IARI, NEW DELHI	Selection from NM92	2002	Resistant to MYMV
8	IPM-99-125 (Meha)	IIPR, KANPUR	PM3 × APM36	2004	Resistant to MYMV
10	HUM-16	BHU, VARANASI	PUSA BOLD-1 × HUM8	2006	Resistant to MYMV
11	MH-2-15	CCSHAU, HISAR	PDM 116 × GUJRAT-1	2007	Resistant to MYMV and Cercosporia leaf spot
12	Pant Moong 6 (UPM02-17)	GBPAU, PANTNAGAR	Pant mung-2 × AMP36	2007	Resistant to MYMV
13	KM-2241	CSAU, KANPUR	SAMRAT × PDM54	2008	Resistant to MYMV
14	IPM-2-3	IIPR, KANPUR	IPM99-125 × PUSA BOLD2	2009	Resistant to MYMV
15	PUSA 0672	IARI, NEW DELHI	11/395 × ML267	2009	Resistant to MYMV
16	IPM-2-14	IIPR, KANPUR	IPM99-125 × PUSA BOLD2	2010	Resistant to MYMV
17	DGGV-2	UAS, DHARWAD	CHINAMUNG × TM-98-50	2014	Moderately resistant to Powdery Mildew
18	MH421	CCSHAU, HISAR	MUSKAN × BDYR2	2014	Resistant to MYMV

19	IPM410-3 (Shikha)	IIPR, KANPUR	IPM03-1×NM1	2016	Resistant to MYMV
20	IPM205-7 (Virat)	IIPR, KANPUR	IPM 02-1 × EC39889	2016	Resistant to MYMV
21	LGG460	APAU, LAM	LAMM2 × ML267	1997	Top bearing with more pod per cluster, Resistant to MYMV
22	Pant Mung 5	GBPAU, PANTNAGAR	Selection from VC 6368	2007	Erect, shining green large seed, t olerant to MYMV
23	TM 96-2	BARC & ANGRAU, LAM	KOPERGAON × TARM 2	2007	Resistant to Powdery Mildew
24	SML-832	PAU, LUDHIANA	SML302 × PUSA BOLD1	2010	Tolerant to MYMV and thrips
25	MH-3-18	CCSHAU, HISAR	ASHA × BDYR 1	2016	Resistant to MYMV
26	SML-668	PAU, LUDHIANA	Selected from NM94	2002	Resistant to MYMV
27	CO -4	TNAU, COIMBATORE	MUTANT OF CO-1	1981	Erect, large seeds suitable for rain fed, resistant to PM & MYM V
28	CO -5	TNAU, COIMBATORE	KM 2 × MG 50-10	1991	Tolerant to MYMV
29	CO -6	TNAU, COIMBATORE	WGG 37 × CO-5	1999	Resistant to MYMV
30	HUM - 1 (Malviya Jyoti)	BHU, VARANASI	BHUM 1 × PANT U30	1999	Resistant to MYMV
31	COGG -912	TNAU, COIMBATORE	MGG 336 × COGG 902	2005	Resistant to MYMV and CLS
32	COGG -8	TNAU, COIMBATORE	COGG 923 × VC 6040	2014	Resistant to YMV
33	PDM 139 (Samrat)	IIPR, KANPUR	ML 20/19 × ML 5	2001	Resistant to YMD
34	IPM 302-2 (Kanika)	IIPR, KANPUR	PANT MUNG 4 × EC 398897	2018	Resistant to MYMV
35	IPM 2K14-9 (Varsha)	IIPR, KANPUR	EC 398885 × PDM 139	2018	Resistant to MYMV and Powder y Mildew
36	IPM 312-20 (Vasudha)	IIPR, KANPUR	IPM 3-1 × SPS 5 Interspecific Hybridization	2020	Moderately resistant to Anthracnose and powdery mildew and resistant to MYM V
37	IPM 409-4 (Heera)	IIPR, KANPUR	PDM2881 × IPM 3- 1	2020	Resistant to MYMV
38	LGG 574	APAU, LAM	Mutant from Pant M2	1993	Resistant to MYMV

39	IPMD-604- 1-7	IIPR, KANPUR	GERMPLASM	-----	ResistanttoMYMV
40	OBBG-58	Ouat, BERHAMPUR	MUTANT OFK857	2002	ResistanttoMYMV

Table 2: ListofPolymorphicSSRMarkers

S.No	Primer Name	Sequences	T _m	Annealing Temperature	Length
1	VR91	R = 3` TGGAGATGCAGGACTAAGAAGAG 5`F=5`ACATATGTATCTGTCTGTGTGCCTA3`	55.3°C 54.4°C	52.4°C	23mer 25mer
2	VR86	R=3`ATCGGTATATGTTGCCAATCAG5`F=5` `CTATACTGCAATGAAGTGGATCTC3`	51.1°C 54.0°C	52.5°C	22mer 24mer
3	VR80	R=3`AATGGTCCCTTTACCCCTTTT5`F=5` `TGTGAGAGTGGAAGAGCAACTT3`	50.5°C 53.0°C	51°C	21mer 22mer
4	VR48	R=3`AATAGGGCCCATAACATGTCC5`F=5` 5`AGGTGAGTGAAAATTGGAATAGG3`	52.4°C 51.7°C	52.5°C	21mer 23mer
5	CEDG 300	R=3`GTGTCGGAAATGTCAGGAGG5`F =5`CGACAAACCCAAACCCTAGC3`	59.4°C 59.4°C	59.4°C	20mer 20mer
6	CEDG 271	R = 3` CACTCCCACTGCCAAACAAGG 5`F=5`GCACTAAAGTTAGACGTGGTTC3`	61.9°C 58.2°C	48°C	21mer 22mer
7	CEDG 236	R=3`CGATGATGAGTCCTTTGGAATTGGG 5` F=5`CTTGGACGGACAGAGTTTGGATTC3`	61.0°C 62.7°C	61.8°C	25mer 24mer
8	CEDG 115	R = 3`ATGCCTCCTTTCAGGTGATTGT 5`F=5`GGCTCATTGTACCACTGGATAT3`	58.4°C 58.4°C	58.4°C	22mer 22mer
9	CEDG 048	R=3`GCTCCTCTTTTTGCTGCATC5`F =5`TCTCTTCTCTATGGCTTGG3`	54.6°C 53.5°C	54°C	20mer 20mer

10	CEDG 021	R=3`AAAGGATGCGAGAGTGTAGC5`F =5`GCAGAATTTTAGCCACCGAG3`	57.8°C 57.8°C	57.8°C	20mer 20mer
----	-------------	--	------------------	--------	----------------

Table3: Grouping of forty forty Greengram genotypes based on D² values

Cluster Group	No. of Genotypes	List of Genotypes
1 Cluster	26	IPM 99-125 (Check), HUM -1, IPM 410-3 (Shikha), MH 2-15, IPM 302-2 (Kanika), DGGV -2, OBGG -58, MH 3-18, PUSA 0672, SML 832, IPM 409-4 (Heera), IPM 205-7 (Virat), IPMD -604-1-7, ADT -3, IPM 2K14-9 (Varsha), PUSA 9072, IPM 2-14, KM 2241, COGG -912, IPM 2-3, LGG 574, HUM 16, MH -421, CO -4, PDM-11 & Pant Mung 5
2 Cluster	7	Pant Mung 4, CO -6, SML 668, TMB -37, PUSA 9531, LGG 460 & PUSA VISHAL
3 Cluster	1	COGG -8
4 Cluster	1	PDM 139 (Samrat)
5 Cluster	3	PDM -54, IPM 312-20 (Vasudha) & TM 96-2
6 Cluster	1	CO -5
7 Cluster	1	Pant Mung 6

Fig. 1: Diagrammatic representation of Cluster distance using Tocher method among 40 Greengram genotypes

UNDER PEER REVIEW

Cluster 1	43.36	95.99	85.36	61.71	300.20	98.52	87.13
Cluster 2		40.76	89.83	181.96	114.23	218.80	188.63
Cluster 3			0.00	180.58	276.84	173.77	180.07
Cluster 4				0.00	432.89	66.62	53.61
Cluster 5					43.67	472.88	424.78
Cluster 6						0.00	50.49
Cluster 7							0.00

UNDER PEER REVIEW

Table 5: Principal Components Analysis for 13 quantitative traits of 40 Greengram genotypes

S. NO		1 Vector	2 Vector	3 Vector	4 Vector	5 Vector	6 Vector
	Eigene Value (Root)	2.94239	2.62338	2.18142	1.40827	1.28463	0.86415
	% Var. Exp.	22.63380	20.17985	16.78012	10.83286	9.88175	6.64734
	Cum. Var. Exp.	22.63380	42.81365	59.59377	70.42663	80.30838	86.95571
1	Days to 50% flowering	0.25407	0.29025	0.40080	0.04205	0.11657	0.27832
2	Days to fifty percent pod	0.00278	0.04107	0.09691	0.04064	0.70550	-0.56331
3	Days to maturity	0.15131	0.37155	0.39661	-0.13677	0.01366	-0.00948
4	Plant height (cm)	0.36226	0.15558	0.16335	0.32086	-0.31572	-0.41856
5	Number of Primary Branches	-0.19349	-0.26814	0.20174	-0.11537	-0.50147	-0.35084
6	Number of clusters per plant	0.30859	-0.03585	-0.31699	0.50778	-0.12445	-0.19631
7	Number of pods per plant	0.28275	0.19573	-0.38377	0.27369	0.02998	0.27913
8	Pod length(cm)	-0.13520	-0.07461	0.48781	0.47262	0.09161	0.16565
9	No. of seeds/pod	-0.10037	-0.37708	0.29164	0.38106	-0.05144	0.16748
10	Seed Index (g)	-0.43522	0.19739	-0.12200	0.29609	0.09212	-0.26275
11	Harvest index (%)	-0.04655	-0.49805	-0.08604	0.11165	0.27589	0.14309
12	Biological yield per plant (g)	0.42710	-0.32967	0.05293	-0.07250	0.15840	0.03493
13	Seed yield per plant (g)	0.40786	-0.31123	0.09707	-0.23432	0.01502	-0.21017

Fig. 2: 3D plot and 2D plot of Principal Components Analysis for 13 quantitative traits of 40 Greengram genotypes

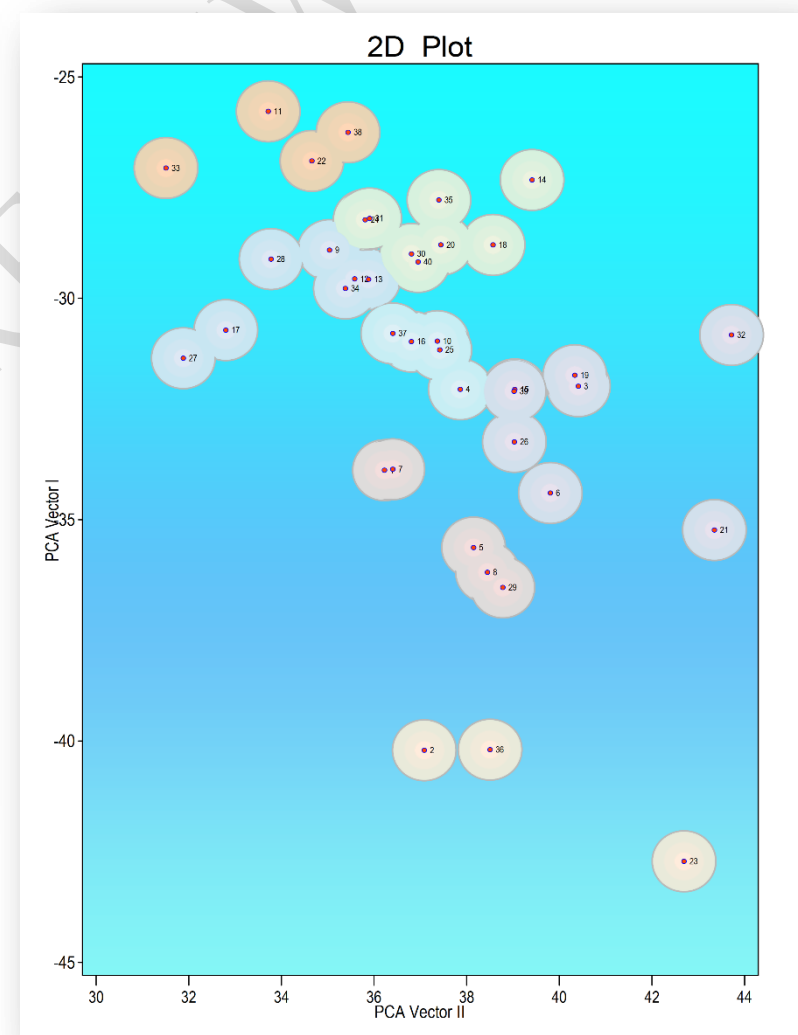
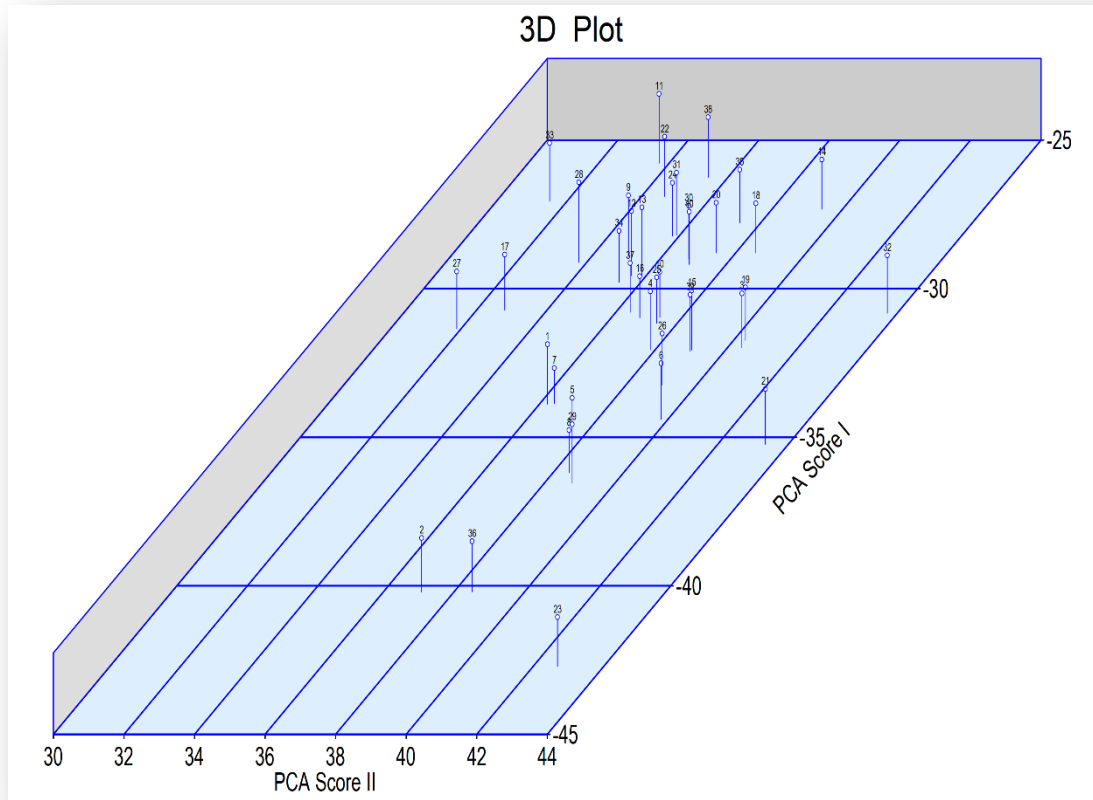


Fig. 3: Distribution of 40 genotypes on the bi-plot axes in PC 1 and PC2 of principal component analysis

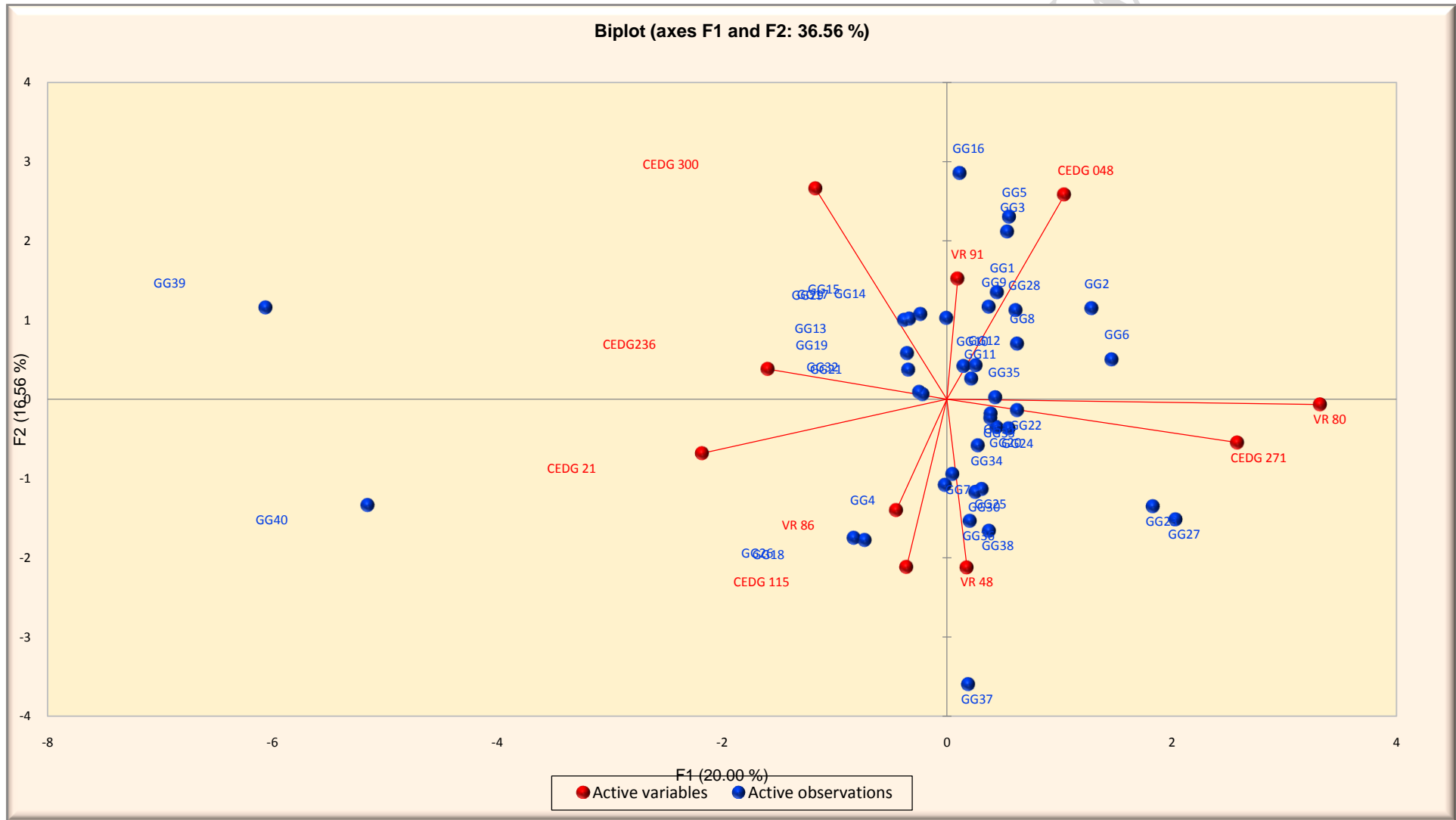


Table6: Total variance explained by different principal components in Greengram

Observations	Eigenvalue	Variability(%)	Cumulative%
PC1	2.000	19.995	19.995
PC2	1.656	16.562	36.557
PC3	1.359	13.595	50.152
PC4	1.293	12.929	63.081
PC5	1.004	10.041	73.122
PC6	0.882	8.824	81.947
PC7	0.685	6.849	88.795
PC8	0.561	5.615	94.410
PC9	0.434	4.343	98.753
PC10	0.125	1.247	100.000

Table 7. PCA analysis of 10 SSR markers for molecular diversity of Greengram

Observations	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
CEDG115	-0.096	-0.516	-0.400	-0.168	0.536	-0.197	0.368	-0.113	0.250	0.012
CEDG300	-0.313	0.647	0.101	0.142	-0.237	-0.158	0.480	-0.374	-0.009	-0.021
CEDG271	0.693	-0.133	-0.184	0.541	-0.219	-0.075	0.208	0.149	0.137	-0.192
CEDG048	0.280	0.629	0.188	-0.460	0.057	0.053	0.195	0.398	0.272	0.046
CEDG236	-0.427	0.093	0.441	0.604	0.190	-0.057	-0.239	-0.018	0.390	0.023
CEDG21	-0.583	-0.166	-0.457	0.311	-0.411	0.015	0.172	0.323	0.010	0.153
VR86	-0.121	-0.341	0.708	0.199	0.256	0.067	0.369	0.226	-0.266	0.007
VR48	0.048	-0.517	0.332	-0.290	-0.417	0.501	0.109	-0.217	0.224	0.046
VR80	0.891	-0.016	0.076	0.283	-0.006	-0.209	0.026	-0.138	-0.021	0.237
VR91	0.026	0.371	-0.315	0.273	0.406	0.712	0.046	-0.077	-0.075	0.051

Table8: Information generated by using 10 SSR Markers in Greengram Genotypes

Markers	Alleles count	Number of effective alleles	Observed heterozygosity	Expected heterozygosity	PIC	I	Fixation index
CEDG115	5	2.896	1.255	0.125	0.655	0.663	0.809
CEDG300	4	2.276	0.972	0.125	0.561	0.569	0.777
CEDG271	2	1.226	0.331	0.000	0.184	0.186	1.000
CEDG236	4	1.237	0.443	0.147	0.192	0.194	0.233
CEDG048	3	1.203	0.363	0.182	0.169	0.172	-0.076
CEDG21	2	1.051	0.117	0.000	0.049	0.049	1.000
VR86	4	3.265	1.258	0.000	0.694	0.703	1.000
VR48	4	2.305	0.952	0.000	0.566	0.573	1.000
VR80	4	2.186	0.886	0.000	0.543	0.549	1.000
VR91	5	2.920	1.215	0.000	0.658	0.666	1.000

Fig. 4: Gel pictures of Agarose stained with Ethidium bromide showing genetic polymorphism among Greengram (*Vigna radiata* (L.) Wilczek) Genotypes using SSR primers.

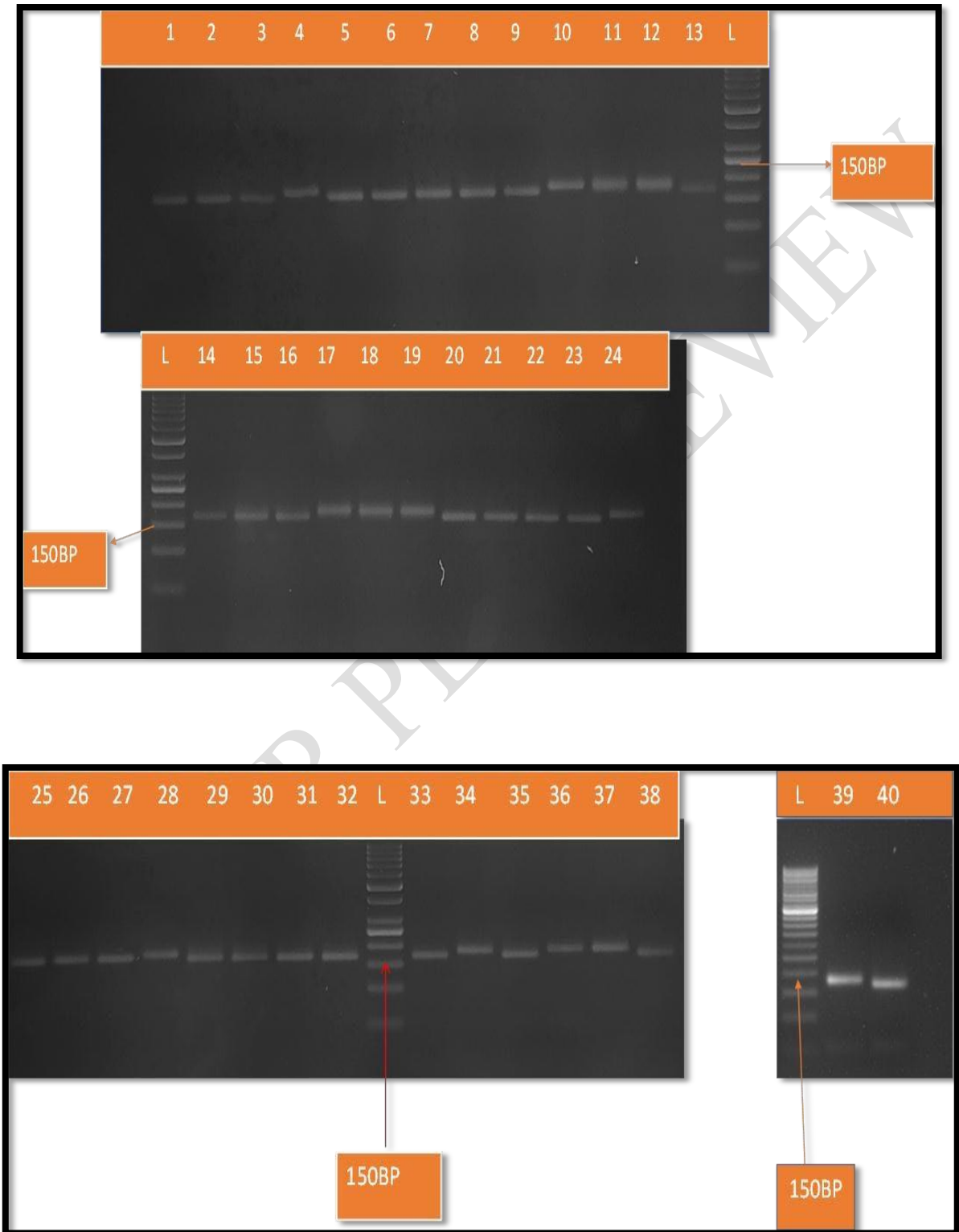
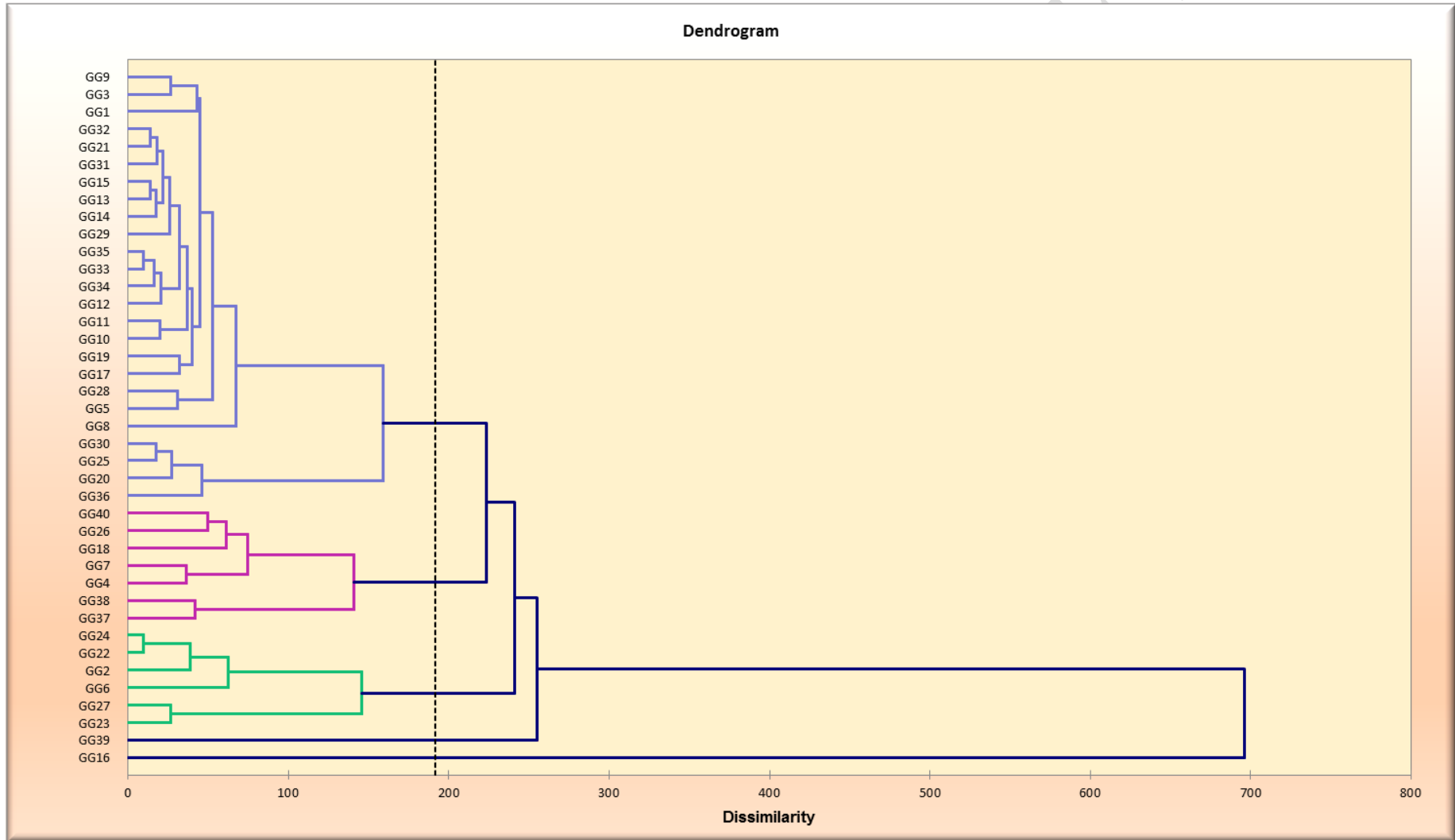


Fig.5: Cluster analysis using ward method and squared Euclidean distance for 40 Greengram genotypes



CONCLUSION

Conclusions can be made based on the findings of the current study, which showed that Analysis of Variance revealed significant differences among the 40 genotypes, suggesting the possibility of selecting promising lines from the available germplasm. Based on average performance, the LGG 574 genotype produced the most seeds, followed by PDM 139 and Pant Mung 6. The separation between clusters 4 and 5 was the maximum, followed by clusters 5 and 6. The genotype compositions of these clusters show a broad range of segregational diversity. 86.95% of the variation for all the attributes was explained by the first six principal components (PC-1 to PC-6), with eigenvalues of 2.94, 2.62, 2.18, 1.40, 1.28, and 0.86, respectively. In VR-86, the PIC and heterozygosity percentages were found to be the greatest. With 25 genotypes, Cluster I have the most genotypes when compared to Clusters II, III, IV, and V. SSRs with significant polymorphism information richness allowed for the differentiation of genotypes in our study. The results of the study show that SSR analysis can be used to evaluate the molecular diversity of various Greengram genotypes. For the purpose of marker-assisted breeding programmes, plant breeders highly value the information generated on marker data.

REFERENCES

- Aijaz, A., Razvi, S.M., Rather, M.A., Ahmad, M., Zaffar, G., Ganie, S.A., Mir, M.R and Hakeem, K.R. (2016) Estimation of genetic divergence in Greengram (*Vigna radiata*(L.) Wilczek) under temperate ecology of Kashmir. *Journal of Botanical Sciences*, 5(1):29-33.
- Changyou, L., Jing, W., Wang, L., Fan, B., Cao, Z., Su, Q., Zhang, Z., Wang, Y., Tian, J., Wang, S. Quantitative trait locus mapping under irrigated and drought treatments based on a novel genetic linkage map in mungbean (*Vigna radiata*(L.) Wilczek). *Theoretical and applied genetics*, 2017:130(11), 2375–2393.
- Doyle, J.J., Doyle, J.L., A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochemical Bulletin*, 1987:19, 11–15.
- Gwag, J.G., Dixit, A., Park, Y.J., Ma, K.H., Kwon, S.J., Cho, G.T., Lee, G.A., Lee, S.Y., Kang, H.K., Lee, S.H. Assessment of genetic diversity and population structure in Greengram. *Journal of Genetics and Genomics*, 2010:32: 299–308.
- Jakhar, N.K. and Kumar, A. Principal component analysis and character association for yield components in Greengram [*Vigna radiata*(L.) Wilczek] genotypes. *Journal of Pharmacognosy and Phytochemistry*, 2018:7(2):3665-3669.
- Kaiser, H.F. The applications of electronic computers to factor analysis. *Educational and Psychological Measurement*, 1960:20:141-151.
- Kanimoli, M., Jagadeeshselvam, N., Madhumitha, B., Karthikeyan, A., Pandiyan, M., Karthikeyan, G., Vanniarajan, C., Raveendran, M., Senthil, N., Sudha, M. Screening and Identification of SSR Markers for Genetic Diversity for mungbean (*Vigna radiata* (L.) Wilczek). *International Journal of Current Microbiology and Applied Sciences* 2018:7(4), 789–793.
- Katiya Manoj and Kumar Sachin. Multivariate analysis for genetic divergence in Greengram (*Vigna radiata* (L.) Wilczek). *International Journal of Science and Research*. 2015:4 (6):462-464.
- Mahalanobis, P.C. On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, 1936:2(1):49-55

- Mahalingam, A., Manivanan, N., Bharathi, K., Ramakrishnan, P. and Vadivel, K. Character association and principal component analysis for seed yield and its contributing characters in Greengram (*Vigna radiata* (L.) Wilczek). *Electronic Journal of Plant Breeding*, 2020:11(1):259-262.
- Muthuswamy, A., Jamunarani, M. and Ramakrishnan, P. Genetic variability, character association and path analysis studies in Greengram (*Vigna radiata* (L.) Wilczek). *International Journal of Current Microbiology and Applied Sciences*, 2019:8(4):1136-1146.
- Ramakrishnan, C. K., Divya, Savithramma, D. L., and Vijayabharathi, A. Studies on Genetic Variability, Correlation and Path Analysis for Yield and Yield Related Traits in Greengram [*Vigna radiata* (L.) Wilczek]. *International Journal of Current Microbiology and Applied Sciences* 2018:7(3):2753-2761.
- Sai Rekha, K.; Reddy, D. M.; Reddy, K. H. P.; Raja Rajeswari, V.; Roopesh Kumar, R. B. and Ravindra, R. Genetic diversity studies under stress condition in mung bean (*Vigna radiata* L. Wilczek), *Electronic Journal of Plant Breeding* 2015:6(1):225-232.
- Sen, M.; De, D. K. Genetic divergence in Greengram. *Legume Research-an international journal*. 2017:40(1):16-21.
- Sharma, S. R.; D. Singh.; Kumar, Pawan.; Khedar, O. P.; Varshnay, N. Assessment of Genetic Diversity in Greengram [*Vigna radiata* L Wilczek] Genotypes. *International Journal of Genetics*. 2018:10(7):471-474.
- Singh, R. K. and B. D. Choudhury. Biometrical methods in Quantitative Genetic Analysis. Kalyani publishers, New Delhi, Ludhiana, India, 1987:318
- Singh, B., Goswami, A. and Kumar, A. 2017. Biotechnology approaches for quality improvement of vegetables. *Annales Horticulture*. 2017:10(1):7-14.
- Suhel, M., Singh, I. P., Abhishek, B. and Singh, Chandra Mohan, S. Multivariate analysis in Greengram (*Vigna radiata* (L.) Wilczek). *Legume Research*. 2015:38(6):758-762.
- Souframanien, J. and Gopalakrishna, T. Cross-

species amplification of microsatellite loci and diversity analyses in blackgram. *Journal of Food Legumes*, 2009; 22(1): 11-17.

Thippani, S., Eshwari, K.B. and Bhave, M.H.B. Principal component analysis for yield components in Greengram accessions (*Vigna radiata* L.). *International Journal of Pure Applied Biosciences* 2017; 5(4): 246-253.

Wang, L., Bai, P., Yuarf, X., Chen, H., Wang, S., Wang, S., Chen, X., Cheng, X. Genetic diversity assessment of a set of introduced mung bean accessions (*Vigna radiata* (L.) Wilczek). *The Crop Journal*. 2018; 6, 207–213.

UNDER PEER REVIEW