RESEARCH ARTICLE

DIVERSITY ANALYSIS IN ADVANCE LINES OF DESI COTTON (*GOSSYPIUM ARBOREUM* L.) USING SSR MARKERS

Kuldeep Jangid, Omender Sangwan and Sagar*

Department of Genetics & Plant Breeding, CCS Haryana Agricultural University, Hisar-125004[1,2,3,]
Email: sagararora470@gmail.com

Received-02.03.2022, Revised-14.03.2022, Accepted-27.03.2022

Abstract: 30 advance lines of desi cotton were used for this study. Total 22 polymorphic markers amplified 62 bands with an average of 2.81 bands per marker. The band fragment size ranged from 120-380 bp. The polymorphism information content (PIC) ranged from 0.234 to 0.767 with an average of 0.518. BNL4108 marker gave maximum polymorphism with PIC value 0.767. HD526 and HD535 are most similar accessions with highest value of similarity matrix, 0.956. NAU2336 revealed least polymorphism with PIC value 0.234 and 2 different allele sizes. Significant diversity was found among all 30 genotypes even at molecular level and all the 30 genotypes were grouped into 4 clustered by polymorphism information of 22 SSR markers.

Keywords: Desi cotton, Diversity, Hisar, NTSYS-PC, SSR markers

INTRODUCTION

Cotton plays an important role among the cash crops from the earliest times. It finds mention in the Rigveda; the oldest scripture of the Hindus. Cotton is popularly called as "White Gold" as it is the most important renewable natural fibre crop and continues to be the predominant and sustainable fibre in the Indian textile industries, despite stiff competition from the man made synthetic fibers. Thus cotton plays an important role in agriculture, industry and economic development of the country. Worldwide area under cotton for the year 2020-21 was 31.66 million hectares and production and productivity accounted for 113.11 million bales and 778 kg/hectare respectively. India has emerged as the largest producer of Cotton in the world and occupies the first position in terms of total area and production. Among the major Cotton exporting countries in the world, India occupied 3rd position with 5.5 million bales (USA–16.25 million bales and Brazil –10.70 million bales). In India during 2020-21 (provisional), production of Cotton was 371.00 lakh bales cultivated under an area of 129.57 lakh hectares with a productivity of 487 kg per hectare (Cotton Outlook Report, Agricultural Market Intelligence Centre, ANGRAU). Cotton is the leading fibre crop of the world which consists of 50 species, out of which 44 are diploid (2n = 2x = 26) and possess A to G and K genomes. The remaining five species are allotetraploids with AD genome (2n = 4x = 52, AADD). The diploid species of A and D genomes are distributed in different hemispheres and show high genetic divergence between them. There are total four cultivated species of cotton among these; diploid species (*G. arboreum* L. and *G. herbaceum* L.) are known as old world cotton and tetraploid (*G. hirsutum* L. and *G. barbadense* L) as new world cotton (Brubaker *et al*. 1999). India is the unique country where all these four cultivated species are grown commercially. In North India only *G. hirsutum* and *G. arboreum* species are grown. *G. arboreum* is mainly grown under poor crop management conditions and their yield potential is not being realized fully. This species possesses special desirable attributes like hardiness, earliness, tolerance to drought and tolerance to insect pests, thereby ensuring low cost of cultivation. Its lint is mainly used for blending purposes and as surgical cotton. Cotton seed is a good source of oil and protein. Cotton seed contains 18-20% seed oil which is edible after removal of gossypol by hydrogenation and 17- 23% seed protein by weight. Cotton productivity and the future of cotton breeding efforts tightly depend on the level of the genetic diversity of cotton gene pools and its effective utilization in cotton breeding programs. The main objective of cotton breeding programs is enhancement of seed cotton yield. Thus, evaluation of genetic diversity related to these characters is important for improvement and sustainable production of cotton. Molecular markers are indispensable tools for improving plant breeding methods by facilitating diversity analysis at seedling stage and early selection through linked genes of interest. These markers are not influenced by the environment and can be scored at all stages of the plant growth, which saves time, resources and energy. Genotype specific markers are very much useful in cotton breeding for varietal registration, protection of PBRs, DNA

*Corresponding Author

fingerprinting and easy detection of useful traits. Molecular markers enhance the efficiency of selection and varietal improvement (Verma *et al*. 2015). SSR are employed in genetic analysis of cotton due to their polymorphic nature (Liu *et al*. 2000, Reddy *et al*. 2001). In genetic analysis of traits, DNA markers are better than morphological and cytological markers because these are (i) highly polymorphic (ii) free from pleiotropic and epistatic effect (iii) enable employment of non destructive methods and require small amount of plant tissue (iv) free from environmental effect and (v) highly reproducible.

MATERIALS AND METHODS

Plant material

Young leaves of the thirty advance lines of desi cotton were collected from Research Farm of Cotton Section, Department of Genetics and Plant Breeding, CCSHAU, Hisar. All these genotypes were developed and grown by Cotton Section (Table 1).

Table 1. List of *Gossypium arboreum* genotypes

S.N	Genotypes.	S.N	Genotypes.
1	HD 324	16	HD 540
2	HD 418	17	HD 541
3	HD 432	18	HD 542
4	HD 503	19	HD 543
5	HD 509	20	HD 544
6	HD 514	21	HD 545
7	HD 521	22	HD 546
8	HD 524	23	HD 547
9	HD 526	24	HD 548
10	HD 534	25	HD 549
11	HD 535	26	HD 550
12	HD 536	27	HD 551
13	HD 537	28	HD 552
14	HD 538	29	HD 553
15	HD 539	30	HD 554

Genomic DNA Isolation

Genomic DNA was isolated from the young leaves of 30 genotypes using CTAB method of Saghai-Maroof *et al*. (1984). Young leaf material, 7-8 leaves were transferred to pre chilled pestle and mortar. Then 10 ml of liquid nitrogen was added and leaf material was ground to fine powder. 15.0 ml quantity of Warm (65°C) CTAB extraction buffer was added (1M tris HCl with pH 8.0, 5M NaCl, 0.5M EDTA with pH8.0, 1.5 ml ß- mercaptoethanol, 2% CTAB and 1% PVP which was prepared fresh. This was added to 50 ml Oakridge tubes containing 10-15 g sample approximately and mixed well by shaking. This was incubated at 65ºC for 1h in the water bath. The cells were thus lysed and denatureation of proteins took place. Tubes were removed from water bath and cooled to room temperature for 5-10 min. Then added 15 ml of chloroform: Isoamyl alcohol (24:1) and mixed by turning gently for 5-10 min followed by centrifugation at 12000 rpm for 15 min. The aqueous layer was poured into new 50 ml Oakridge tubes and 10 ml chloroform: Isoamyl alcohol (24:1) was added and centrifuged at 10000 rpm for 15 min. Now the upper aqueous layer was transferred in 50 ml Oakridge tubes and then 10 ml of chilled isopropanol was added and mixed by gentle inversion to precipitate DNA or leave it overnight at -20°C for precipitation. DNA was taken out and put into 1.5 ml centrifuge tubes. DNA pellet was made by centrifugation at 8000 rpm. The supernatant was poured off and DNA pellet was washed with 70% ethanol twice by leaving for 20 min. Ethanol was then poured off and the DNA pellet was dried at room temperature for 30 minutes. Ethanol washing helps in the removal of salts. The pellet was dissolved in 150 µl of TE. The dissolved DNA was stored at 20°C for further use.

Estimation of quality and quantity of DNA

The DNA concentration was checked by visual assessment of band intensity in comparison with

known concentration in 0.8% agarose gel **(Figure 1).** A single discrete band near the wells showed that genomic DNA was intact and free from any contaminations. The final genomic DNA concentration was diluted to 50-100 ng. Forty SSR markers were used to screen the genotypes out of those twenty two polymorphic SSR primers were used to generate DNA profile of 30 genotypes. **(Table 2)**

Agarose gel electrophoresis and allele scoring

PCR amplified DNA fragments for SSR markers were resolved by submerged horizontal electrophoresis in 1 % (w/v) Agarose gels. Polymorphic bands were scored based on the presence or absence of bands for each genotype. Scoring of clearly visible bands was done.

Data analysis

Polymorphism frequency was calculated on the basis of presence or absence of common bands between selected lines for each marker. A genetic similarity coefficient matrix based on the Jaccard's coefficient was computed for the construction of phylogenic tree using NTSYS-PC to estimate the genetic similarities among the genotypes.

Table 2. List of Polymorphic markers with their chromosome numbers, sequence and Annealing temperature.

Sr. No.	Marker Name	Ch. No.	Forward primer (5' to 3')	Reverse primer (5' to 3')	Annealing Temp. (°C)
1	NAU1070	3	CCCTCCATAACCAAAAGTTG	ACCAACAATGGTGACCTCTT	58.90
2	NAU1093	6	TGTGATGAAGAACCCTCTCA	AAATGGCGTGCTTGAAATAC	59.20
3	NAU1103	21	GGAGCCAGAAGTTGAGAAAA	TTCGGCTTCTGCTTTTACTT	58.80
4	NAU1141	13	CCCCTCTCTCTGTTTCTCAA	AAGGGGTTTGAAGGGTTATC	58.80
5	NAU1156	5	ACACTCTCTCAGCTGGAACC	GGTCTCCCTCTAGCTTGTTG	58.80
6	NAU1167	3	CTGACTTGGACCGAGAACTT	AAGAGCCCTGGACAATGATA	59.00
7	NAU1233	10	TTCGGGAAAGTTAGAGGAGA	TCCTCAGAGCTCGGAATAGT	58.50
8	NAU1369	21	TGGCAGAGATGAATGTAAGC	GGTAACGGATGGAAAATCAC	58.50
9	NAU2083	1	AGAAGAGGTTGACGGTGAAG	TGAGTGAAGAACCTGCACAT	59.00
10	NAU2162	17	ACACAAAAACCCAAAGGAAA	CACGAGTGTCCTTGCTACAG	59.00
11	NAU2238	25	TTTTTCATGGCTGAACTTTG	ATTTCCATTGCTCGCTTTAC	58.50
12	NAU2265	2	CAATCACATTGATGCCAACT	CGGTTAAGCTTCCAGACATT	58.90
13	NAU2336	14	TGGAAAAGGAAGAGGAGAGA	CCCTGAAGTTGTCAAGCTCT	58.80
14	NAU2355		ACAAACAAAACGCCTTCTTC	AACACAAAAACGGTTCCAGT	58.80
15	NAU2697	26	ATTTCCCATGGTCATAGCAG	GAAAGGAGTCGGAAATGAGA	59.10
16	BNL1053	21	AGGGTCTGTCATGGTTGGAG	CATGCATGCGTACGTGTGTA	62.20
17	BNL1672	6	TGGATTTGTCCCTCTGTGTG	AACCAACTTTTCCAACACCG	61.70
18	BNL3089	4	TCGAACTTAACAAAAGAGAGAGAGA	AAATCCGGATTCAGCAATACTT	59.90
19	BNL3649	7	GCAAAAACGAGTTGACCCAT	CCTGGTTTTCAAGCCTGTTC	61.70
20	BNL4108	6	TCCACCATTCCCGTAAATGT	TGGCCAAGTCATTAGGCTTT	61.70
21	JESPR220	17	CGAGGAAGAAATGAGGTTGG	CTAAGAACCAACATGTGAGACC	59.80
22	JESPR274	9	GCCCACTCTTTCTTCAACAC	TGATGTCATGTGCCTTGC	59.70

Ch-Chromosome, Tm- Teperature

RESULTS AND DISCUSSION

Total 22 polymorphic markers amplified 62 bands with an average of 2.81 bands per marker. The band fragment size ranged from 120-380 bp. The polymorphism information content (PIC) ranged from 0.234 to 0.767 with an average of 0.518. BNL4108 marker gave maximum polymorphism with PIC value 0.767 and 5 different allele sizes (Figure. 2) followed by marker NAU2355 with PIC value 0.730 and 3 different allele sizes. NAU2336 revealed least polymorphism with PIC value 0.234 and 2 different allele sizes (Table 3). Similarity matrix was obtained on the basis of molecular data using NTSYS-PC software. It revealed that HD526 and HD535 are most similar accessions with highest value of similarity matrix, 0.956, further accessions HD540 and HD542 with similarity matrix value 0.942 observed to be very closely related. However, maximum distantly related genotypes are HD324 to HD538 and HD432 to HD551 with similarity matrix value 0.347. Dendrogram obtained from NTSYS- PC software (2.1) using binary data (0/1) of 22 markers, all thirty genotypes were discriminated successfully (Figure. 3). The genotypes were classified into four groups at level of GS= 0.62. Group 2 is considered as

major group which include 16 accessions and Group 1, 3 and 4 are minor with 4 genotypes each. While HD549 and HD550 did not grouped in any cluster and were totally separate. At the level of GS=0.73 major groups further separated into three sub-groups with 4, 3 and 9 genotypes.

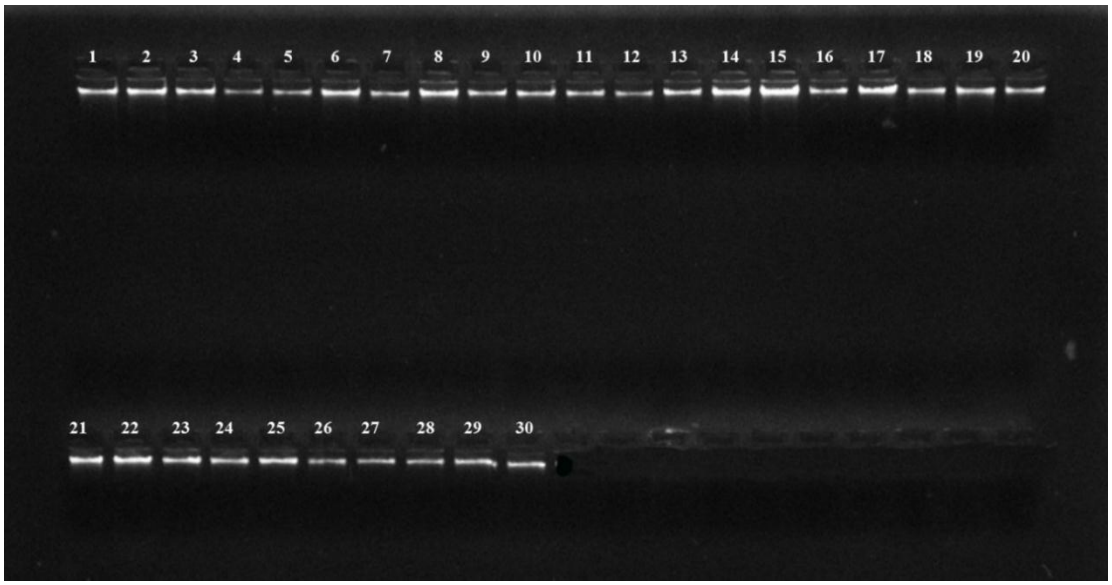

Figure 1: Quality evaluation of isolated DNA of all 30 genotypes on 0.8% agarose gel.

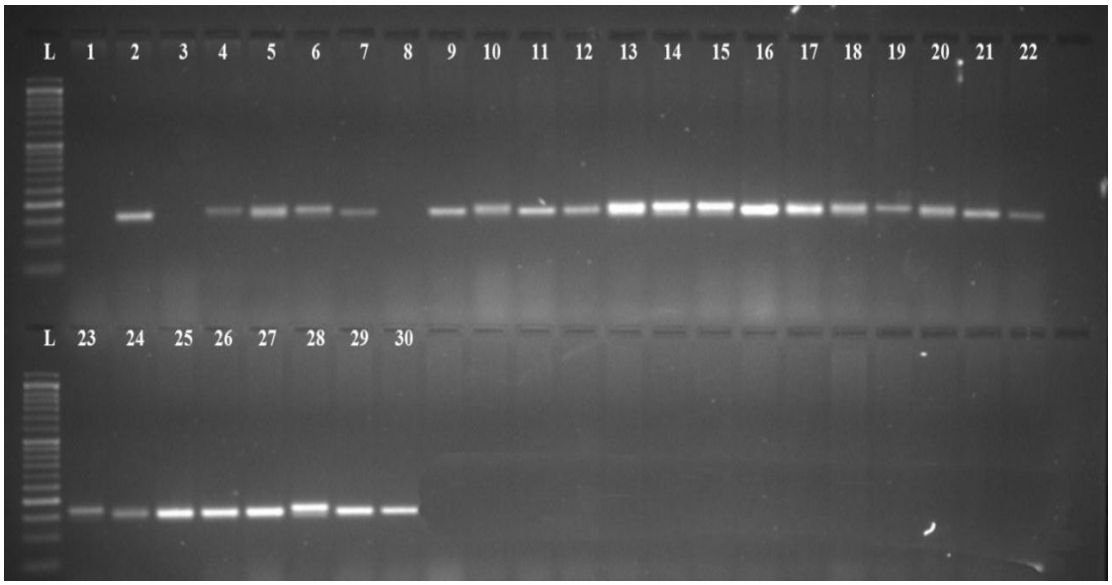

Figure 2: Gel electrophoresis profile of 30 cotton genotypes using primer BNL 4108. The number 1-30 indicates the genotypes and (L) represents 100bp ladder

Table 3. Details of amplified alleles rang and PIC of all polymorphic markers

Sr.No.	SSR locus	Amp. Range (bp)	Number of alleles	PIC
1	NAU1070	150-210	3	0.418
2	NAU1093	150-180	4	0.407
3	NAU1103	200-250	2	0.358
4	NAU1141	220-320	3	0.658
5	NAU1156	210-300	3	0.604
6	NAU1167	190-300	2	0.422

7	NAU1233	260-380	2	0.556
8	NAU1369	250-320	2	0.434
9	NAU2083	160-270	3	0.550
10	NAU2162	200-260	2	0.569
11	NAU2238	120-170	3	0.597
12	NAU2265	150-180	3	0.721
13	**NAU2336**	**130-160**	**2**	**0.234**
14	NAU2355	150-220	3	0.730
15	NAU2697	190-300	3	0.417
16	BNL1053	180-280	4	0.589
17	BNL1672	165-250	3	0.501
18	BNL3089	180-290	3	0.279
19	BNL3649	180-270	3	0.541
20	**BNL4108**	**235-270**	**5**	**0.767**
21	JESPR220	120-300	2	0.498
22	JESPR274	160-200	2	0.562

PIC-Polymorphism Information Content

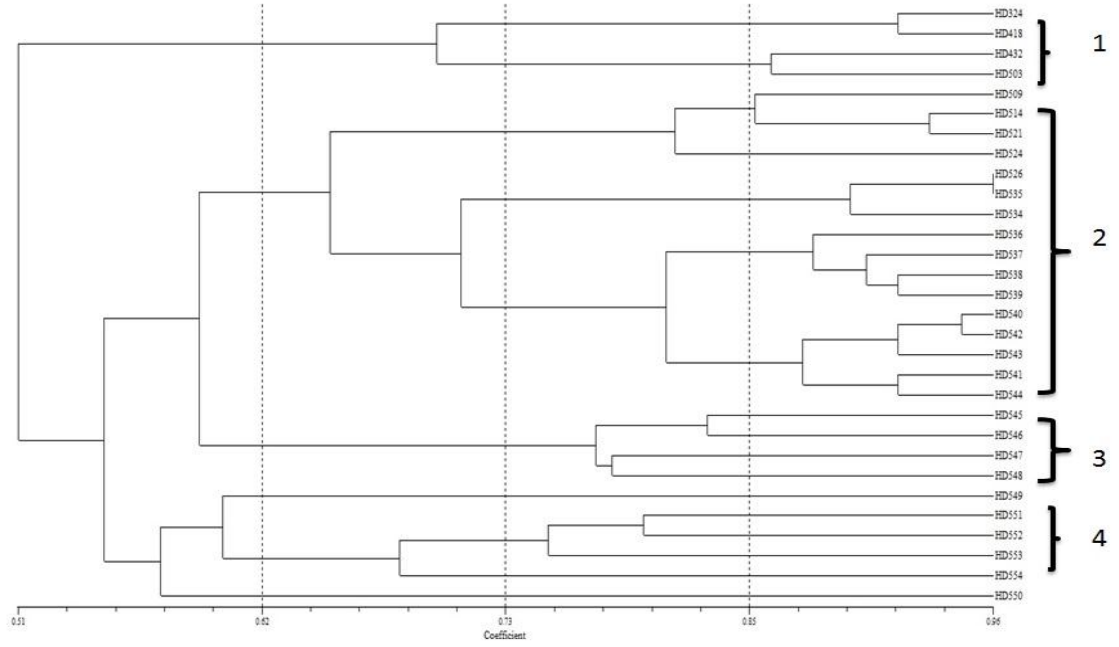

Figure 3. Dendrogram showing relationship among 30 genotypes generated by UPGMA analysis based on polymorphic SSR markers

Forty SSR markers were used to screen the genotypes out of those twenty two polymorphic SSR primers were used to generate DNA profile of 30 genotypes. Total 22 polymorphic markers amplified 62 bands with an average of 2.81 bands per marker. The band fragment size ranged from 120-380 bp. The polymorphism information content (PIC) ranged from 0.234 to 0.767 with an average of 0.518. BNL4108 marker gave maximum polymorphism with PIC value 0.767 and 5 different allele sizes followed by marker NAU2355 with PIC value 0.730 and 3 different allele sizes. NAU2336 revealed least polymorphism with PIC value 0.234 and 2 different allele sizes. Similar results were found by Ijaz *et al*. (2013), Moiana *et al*., (2015), Kavithamani and Amalabalu (2017), Kenchareddi *et al*. (2018) and sagar *et al*. (2020)

CONCLUSION

Forty SSR markers were used to screen the genotypes out of those twenty two polymorphic SSR primers were used to generate DNA profile of 30 genotypes. Total 22 polymorphic markers amplified 62 bands with an average of 2.81 bands per marker. Molecular markers are great tool for genetic diversity analysis, genome mapping, association mapping, marker assisted selection, QTL mapping and for gene pyramiding. Hence on the basis of this study's outcome we can go for best heterotic combination using the most diverse lines.

REFERENCES

Brubaker, C. L., Paterson, A. H. and Wendel, J. F. (1999). Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome, 42: 184-203.

Google Scholar

Cotton Outlook Report. (2021) Agricultural Market Intelligence Centre, ANGRAU, 1-7.

Ijaz Ambreen, Sadia Ali, Usman Ijaz, Smiullah and Tayyaba Shaheen. (2013). Molecular characterization of cotton using simple sequence repeat (SSR) markers and application of genetic analysis. International Journal of Genetics and Molecular Biology, **5** (4): 49-53.

Google Scholar

Kavithamani, D. and AmalaBalu, P. (2017). Studies of genetic parameters on seed cotton yield and fibre quality traits in Egyptian cotton (*Gossypium barbadence*L.). Journal of Cotton Research and Development, **31(2):** 186-193.

Google Scholar

Kencharaddi, H.G., Hanchinal, R.R., Patil, S.S., Manjula, S.M., Kulkarni, Ravi, Usharani, C.V. and Soregoan, C.D. (2018). Molecular Characterization of Heterotic Groups of Cotton through SSR Markers.International Journal of Current Microbiology and Applied Sciences, **7**(03): 426-434.

Khan, S. A., Hussein, D., Askaria, E., Stewart, J., Matik, K. A. and Zafer, Y. (2000). Molecular phylogeny of Gossypium species by DNA fingerprinting. Theor Appl Genet., 101:931-938.

Google Scholar

Liu, S., Cantrell, R. G. and McCarty, **J.C.** (2000). Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. Crop Science, 40:1459-1469.

Google Scholar

Moiana, L.D., Filho, P.S.V., Gonçalves-Vidigal, M.C and Carvalho, L.P. (2015). Genetic diversity and population structure of upland cotton Brazilian cultivars (*Gossypium hirsutum*L. race *latifolium*H.) using SSR markers. African Journal of Crop Science,**9**(2): 143-152.

Google Scholar

Park, Y.H., Alassbady, M. S., Ulloa, M., Sickler, B., Wilkins, T. A., Yu., D. M., Stelly, J., Kohel, R. J., El-Shihy, O. M. and Cantrell, R.G. (2005). Genetic mapping of new cotton fibre loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. Mol Gen Genomics, 274:428-441.

Google Scholar

Reddy, O. U. K., Pepper, A.E., Abdurakhmonov, I., Saha, S., Jenkins, J. N., Brooks, T., Bolek, Y. and El-Zik, K.M. (2001). New dinucleotide and trinucleotide microsatellite marker resources for cotton genome Research. Journal of Cotton Sci, 5:103-113.

Google Scholar

Saghai-Maroof, MA, Soliman, KM, Jorgensen, RA, and Allard, RW. (1984). Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proceedings of the National Academy of Sciences of the United States of America 81(24):8014-8019.

Google Scholar

Sagar, Nimbal, S., Jangid, K. and Kumar, P. (2020). Molecular diversity analysis among elite genotypes of American cotton (*Gossypium hirsutum* L.) Pantnagar Journal of Research, 18(1):1-4.

Google Scholar

Verma, S.K., Dhanda, M. and Salar, R.K. (2015). Analysis of genetic diversity among Asiatic cotton (*Gossypium arboreum* L.) cultivars and breeding lines using RAPD and SSR markers. Int. J. Adv. Res. Biol.Sci. 2(3): 114–122.