

MOLECULAR PROFILING OF CHILLI GERMPLASM BY USING SSR MARKER

M.I. Haque¹, S. Ishtiaque², M.M. Islam^{3*}, T.A. Mujahidi⁴ and M.A. Rahim⁵

¹Regional Spices Research Centre, ²On-Farm Research Division, ³Plant Pathology Division, ⁴Plant Breeding Division

Bangladesh Agricultural Research Institute, Gazipur

⁵Department of Horticulture, Bangladesh Agricultural University, Mymensingh

ABSTRACT

The molecular characterization of chilli germplasm was done based on estimation of genetic diversity among the germplasm by using SSR markers. Forty chilli germplasms were analyzed using eight SSR primers. The SSR primers produced 30 SSR loci with an average value of 3.75 alleles per SSR locus. The similarity index matrix ranged from zero to 2.74. Polymorphic information content (PIC) of the SSR primers ranged from 0.543 to 0.735 with an average value of 0.658. The highest number (five) of allele was observed in primer CAMS-647, whereas the primers CAMS-864, CAMS-880 and CAMS-885 showed lowest number (three) of allele. The smallest allele was found in case of primer CAMS-236 (176 bp), while the longest allele was detected for the primer CAMS-864 (288 bp). Based on similarity matrix using the un-weighted Pair Group Method of Arithmetic Means (UPGMA) dendrogram, chilli germplasms were grouped into four main clusters. SSR markers showed genetic variability in the studied chilli germplasm.

Keywords: Chilli germplasm, Molecular profiling, SSR marker.

INTRODUCTION

Chilli (*Capsicum spp.*), member of Solanaceae family, is one of the most important spice crops of Bangladesh with immense economic significance. In fact, the particular spice crop is originated from South and Central America, where it is still under cultivation (Pickersgill, 1997). In Bangladesh, it is cultivated both in Rabi and Kharif season. Total production of chilli in Bangladesh was approximately 1.30 Lac Metric Tons from 1.02 Lac hectares of land (BBS, 2016). Molecular markers are powerful tools in complementing phenotypic characterization in detecting additional

* Corresponding author: monirulbari@yahoo.com

sources of genetic diversity present within the gene pool. Molecular markers such as isozymes, Restriction Fragment Length Polymorphisms (RFLP), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphisms (AFLP) and Simple Sequence Repeats (SSR) have been used in studying the genetic diversity in *Capsicum spp.* (Tam et al., 2005). These molecular markers are large in number and useful in determining genetic variability through the construction of linkage maps (Gupta et al., 1996). Of the molecular markers developed, SSR markers stand out as exceptional in genetic diversity studies, because they are highly polymorphic and widely distributed in the chilli genome (Mimura et al., 2012). SSR markers, being co-dominant, are able to distinguish genetic relationships between genotypes based on specific traits and are more effective for inbred lines and breeding materials with special attributes (Tam et al., 2005). The extent of genetic variability within a species is vital for its continued existence and adaptation in different agro-ecologies. The more diverse the population is the better for the breeder in developing elite cultivars through careful selection of superior parents. Therefore, an understanding of the genetic variability of a population, through the use of molecular markers, is of critical importance in developing effective strategies for germplasm conservation and breeding purposes (Se-Jong et al., 2012). In light of the frugal information available in Bangladesh with reference to the molecular characterization of chilli germplasm, the current research made an endeavor to characterize applying SSR technique.

MATERIALS AND METHODS

The seedlings of chilli were grown at Regional Spices Research Centre (RSRC), Bangladesh Agricultural Research Institute (BARI), Gazipur, Bangladesh. All nursery/agronomical management practices were done as per standard protocol. The experimental plots were manured and fertilized properly. Thirty-six chilli germplasms were collected from different areas of the country including Spices Research Centre, BARI, Gazipur, of which twenty were indigenous and fifteen were exotic (Table 1). The molecular characterization of the germplasms using simple sequence repeat (SSR) primers to detect polymorphism between the germplasm was carried out at the Biotechnology Laboratory of PGRC at BARI.

Table 1. Germplasm codes with sources of collection for chilli

SL. No.	Germplasm code	Source
1.	CO 001	
2.	CO 002	
3.	CO 003	
4.	CO 446	
5.	CO 446-1	
6.	CO 525	
7.	CO 525-1	SRC, BARI
8.	CO 525-2	
9.	CO 525-3	
10.	CO 610	
11.	CO 610-1	
12.	CO 611-1	
13.	CO 611-2	
14.	CO 613	
15.	CO 626	
16.	CO 629	
17.	CO 630	
18.	CO 631	
19.	CO 632	
20.	CO 633	
21.	CO 634	
22.	CO 635	The World Vegetable Centre
23.	CO 636	
24.	CO 637	
25.	CO 638	
26.	CO 639	
27.	CO 640	
28.	CO 641	
29.	CO 642	
30.	CO 643	
31.	CO 644	
32.	CO 645	
33.	CO 646	
34.	CO 647	SRC, BARI
35.	CO 648	
36.	BARI Chilli-2	

Genomic DNA extraction

Fully opened middle aged leaves were taken from the seedlings into an ice chest and sent to the laboratory for DNA extraction. Genomic DNA was extracted from the 36-chilli germplasm using CTAB method described by Dellaporta et al (1983). Extracted DNA of each accession was confirmed by electrophoresis on 1.0% agarose gel.

SSR (microsatellite) markers and PCR amplification

Simple sequence repeats primers (SSR) used to detect polymorphisms among the chilli germplasm are presented in Table 2. These eight SSR primers are highly polymorphic and widely distributed in the chilli genome (Mimura et al., 2012). They were procured from Metabion International AG (Germany). PCR reactions were carried out in a Techne Thermalcycler (TC-412) in a 10 µl reaction mixture in 96-well plates. PCR kits (KAPA 2G Fast ReadyMix with dye) procured from KAPA Biosystems Ltd (South Africa) was used for the amplification. The kits composed of 2 X PCR master mix containing KAPA2G Fast DNA Polymerase (0.2 U per 10 µl reaction), KAPA2 Fast PCR buffer, dNTPs (0.2 mM each at 1X), MgCl₂ (1.5 mM at 1X), stabilizers and loading dye. 1 µl genomic DNA and 0.5 µl each of forward and reverse primers were added to the PCR kits for DNA amplification

Table 2. SSR primers code and their sequences

SL. No.	Primer code	Forward Sequence	Reverse Sequence
1	CAMS-065	CCAGTCTCATCCAGCAGACA	CATATGCTGCTCCTGCATTC
2	CAMS-075	ACTAATTACACATTCTGCATTTTCTC	AGGCTCGAGTACCACGAAGA
3	CAMS-236	TTGTAGTTTGCGTACCATTGA	ATGAATCCAGGGTTCCACAA
4	CAMS-647	CGGATTCGGTTGAGTCGATA	GTGCTTTGGTTCGGTCTTTC
5	CAMS-855	AAGTGTC AAGGAAGGGGACA	CCTAACCACCCCAAAAAGTT
6	CAMS-864	CTGTTGTGGAAGAAGAGGACA	GCTTCTTTTCAACCTCCTCCT
7	CAMS-880	GAGCCAAGAAAAAGGTGGAA	CAACTCATCGTTCAACAACACA
8	CAMS-885	AACGAAAAACAAACCCAATCA	TTGAAATTGCTGAAACTCTGAA

PCR was subjected to initial denaturation at 95°C for 3 min, followed by cycles of 95°C for 10 sec, 52°C for 10 sec and 72°C for 10 sec. The reaction was repeated for 35 cycles and a final extension at 72°C for 10 minutes was carried out. The reactions were then held at 4°C until electrophoresis. DNA bands were scored as either present (1) or absent (0) for each of the germplasm by visual inspection. Bands with clear and good characteristics were considered and recorded. Loci were considered polymorphic if more than one allele was detected. Cluster analysis of the molecular

data was carried out using NTSYS statistical package (version 4) to generate dendrograms based on genetic similarity matrix using unweighted pair group method with arithmetic mean (UPGMA) method. Power Marker version 3.25 (Liu and Muse, 2005) was used to generate allele frequency, allele number, gene diversity, heterozygosity and polymorphism information content (PIC).

RESULTS AND DISCUSSION

Allelic and loci variation within chilli germplasms

A total of 8 microsatellite primers were utilized to provide genetic diversity among 40 chilli germplasms including 4 stock DNA samples. Among eight primers, six showed polymorphism, CAMS-075 and CAMS-236 was monomorphic. A total of 30 alleles were amplified in 40 germplasms. The number of alleles ranged from 3 - 5 per locus. The average number of alleles is 3.75. The highest number (five) of allele was observed in primers CAMS-647, whereas three primers CAMS-864, CAMS-880 and CAMS-885 showed lowest number (three) of allele. The level of polymorphisms among the 40 chilli germplasms was evaluated by calculating allele numbers and PIC values for each of the 8 SSR loci (Table 3). As a measure of the informativeness of microsatellite, the average PIC value is 0.658 with the range of 0.543 (CAMS -885) to 0.735 (CAMS -647). Size and frequency of all the 30 alleles at 8 microsatellite loci, have been shown in the Table 4.

Genotypic performance of 40 chilli germplasms

Six microsatellite primers exhibited genetic diversity among 40 chili germplasms. The primers exhibited polymorphism in all experimental germplasms. The values of pair-wise comparisons of Nei's (1972) genetic distance between the germplasms were computed from combined data for the eight primers, ranged from zero to 2.74. Comparatively higher genetic distance was observed between CO-611 vs. CO-525-2, CO-638 vs. CO-613, CO-648 vs. CO-638 and CO-003 vs. CO-638. Genetic distance between germplasm can be used to evaluate the genetic diversity of different germplasm (Table 3). The observed number of alleles ranged from three (CAMS-864 CAMS-880, CAMS-885) to a maximum of five (CAMS-647), with an average of 3.75. Overall, the effective number of alleles was less than observed number of alleles, ranging from 2.19 (CAMS-885) to 3.63 (CAMS-236) with a mean of 3.013. Gene diversity expressed by Shannon's information index (I) ranged from 0.92 at CAMS-885 to 1.42 at CAMS-647, with a mean of 1.175. Values for Wright F_{st} showed lowest 0.93(CAMS-065) to highest 1.0 (CAMS-864, CAMS-236, CAMS-647, CAMS-855), with an average of 0.97. Data also revealed lower rate of gene flow with an average of 0.007 (Table 5). The observed heterozygosity varied between

zero (CAMS-647, CAMS-855) and 0.108(CCAMS-065), with a mean of 0.04. Whereas expected heterozygosity varied between 0.55(CAMS-885) and 0.74(CAMS-647), with a mean of 0.66 and Nei's (1973) expected heterozygosity varied between 0.54 (CAMS-885) and 0.73 (CAMS-647), with a mean of 0.66. Overall observed heterozygosity was lower than the expected heterozygosity (Table 6).

Forty chilli germplasm were used to make dendrogram based on similarity coefficient using Unweighted Pair Group Method of Arithmetic Means (UPGMA). Based on 2.055 similarity of co-efficient index, all the 40 chilli germplasm have been differentiated into 4 main clusters: Germplasm 626, 633 and 645 (Net) were grouped in cluster 1 (Fig.1). CO-003, CO629, CO525-3, CO-002, CO613, CO-001, CO647, CO640(Net), CO610-1, CO648 (Blue), CO 641 and CO525-2 (Net) were grouped in cluster 2; CO-629, CO-638, CO632 (Net), CO646, CO525, CO611-2, CO611-1, CO635, CO611, CO631, CO610 (Net), CO637 (Net), CO525 (2), CO611 (1), CO646 (1) were grouped in cluster 3 and germplasm CO644 (Net), BARI Chili-2, CO642 (Net), CO636 (Net), CO634 (Net), CO643 (Net), CO635 (Net), CO525-1, CO525 (Net) and CO446 were grouped in cluster 4.

The SSR analysis showed that the polymorphic level in this study considerably high (75%) compared to earlier reports indicating high level of genetic diversity among the germplasm. For example, Akatas et al. (2009) using 4 primers detected 26% polymorphism in Turkish genotypes; Kochieva and Ryzhova (2005) used 9 primers that showed 8.03% polymorphism. The observed number of alleles in this study ranged from 3 to 5 and the average number of alleles per locus is 3.75, which indicate a greater magnitude of diversity among the germplasm. This is in consistence with the earlier findings (Tilahun et al., 2013). PIC value is the indicator that measures the ability of a marker by considering both number of alleles at a locus and relative frequencies of these alleles. This value depends on the genetic diversity among the population. Markers with PIC value of 0.5 or above are highly informative for genetic studies and are extremely useful in determining the polymorphism rate of a marker at a specific locus (DeWoody et al., 1995). Current study showed PIC value ranged from 0.54 to 0.73, with an average of 0.66. The highest PIC value found was 0.735 for CAMS-647, indicating its most informative and suitable marker among them.

Higher genetic diversity is useful in breeding program to obtain potential genetic gain. In this study genetic distance value ranged from zero to 0.74, indicating germplasm were derived from different origin and could be utilized in breeding purpose for desirable traits. From the contrast between highest and lowest value, it could be said that there were wide variabilities among 40 chilli germplasms. High

Genotypes	446	446(1)	525	525(Net)	525-1	525-2	525-2(Net)	525-3	610(Net)	610-7	611	611-	611(1)	611-1	611-2	613	626	629	631	632(Net)
Pa.Monpura	1.980	2.047	0.762	1.354	0.439	0.661	0.949	0.794	0.949	2.047	0.109	****								
Monpura-1	1.253	1.607	1.575	1.096	0.587	1.320	1.607	1.607	1.096	2.013	1.253	0.882	****							
647-																				
Monpura	1.320	1.386	0.794	1.386	0.877	0.000	0.981	0.693	0.981	2.079	0.626	0.794	0.914	****						
648 (Blue)	0.403	1.386	1.131	2.079	1.570	1.386	1.386	0.981	1.386	0.981	2.013	2.740	1.320	0.693	****					
CO-629	1.320	1.386	1.642	0.000	1.059	0.981	2.079	1.386	2.079	2.079	0.403	0.342	0.914	0.693	1.386	****				
CO-0001	2.013	0.693	1.642	1.386	1.976	0.000	2.079	1.386	2.079	2.079	1.320	1.642	1.320	0.693	0.693	0.981	****			
CO-0003	0.914	1.386	1.642	2.079	1.976	1.386	2.079	1.386	2.079	0.981	2.013	2.740	1.320	0.981	0.288	2.079	0.470	****		
BARI																				
Morich-2	0.847	1.320	2.674	0.000	2.602	1.320	0.000	2.013	0.000	0.914	1.253	1.575	2.639	0.626	0.626	1.320	0.914	0.626	****	
BARI																				
Morich-3	2.013	1.386	1.131	1.386	1.059	2.079	1.386	0.981	1.386	1.386	1.320	1.642	2.013	0.470	0.981	0.981	0.470	0.693	0.914	****

Table 4. Size and frequencies of alleles and diversity index at eight SSR loci across forty chilli germplasm

Locus	No. of Allele	Allele sizes (bp)	Allele frequency	PIC
CAMS-075	4	216	0.154	0.702
		204	0.372	
		194	0.346	
		184	0.128	
CAMS-864	3	288	0.350	0.605
		251	0.500	
		236	0.150	
CAMS-880	3	225	0.250	0.632
		213	0.486	
		202	0.264	
CAMS-065	4	227	0.054	0.634
		225	0.108	
		212	0.405	
		200	0.432	
CAMS-236	4	206	0.180	0.725
		199	0.205	
		187	0.231	
		176	0.385	

Locus	No. of Allele	Allele sizes (bp)	Allele frequency	PIC
CAMS-885	3	239	0.138	0.543
		228	0.613	
		217	0.250	
CAMS-647	5	244	0.075	0.735
		229	0.325	
		219	0.275	
		211	0.050	
		189	0.275	
CAMS-855	4	287	0.200	0.685
		252	0.250	
		239	0.450	
		230	0.100	

Table 5. Summary of genetic variation statistics for all loci

Locus	*na	*ne	*I	*Fst	*Nm
CAMS-075	4	3.354	1.286	0.965	0.009
CAMS-864	3	2.532	0.999	1.000	0.000
CAMS-880	3	2.714	1.049	0.947	0.014
CAMS-065	4	2.733	1.127	0.927	0.020
CAMS-236	4	3.630	1.339	1.000	0.000
CAMS-885	3	2.190	0.920	0.931	0.019
CAMS-647	5	3.774	1.419	1.000	0.000
CAMS-855	4	3.175	1.258	1.000	0.000
Mean	3.750	3.013	1.175	0.973	0.007
St. Dev	0.707	0.558	0.178		

*na = Observed number of alleles, ne = Effective number of alleles, I = Shannon's Information Index and Nm = Gene flow estimated from $F_{st} = 0.25(1 - F_{st})/F_{st}$.

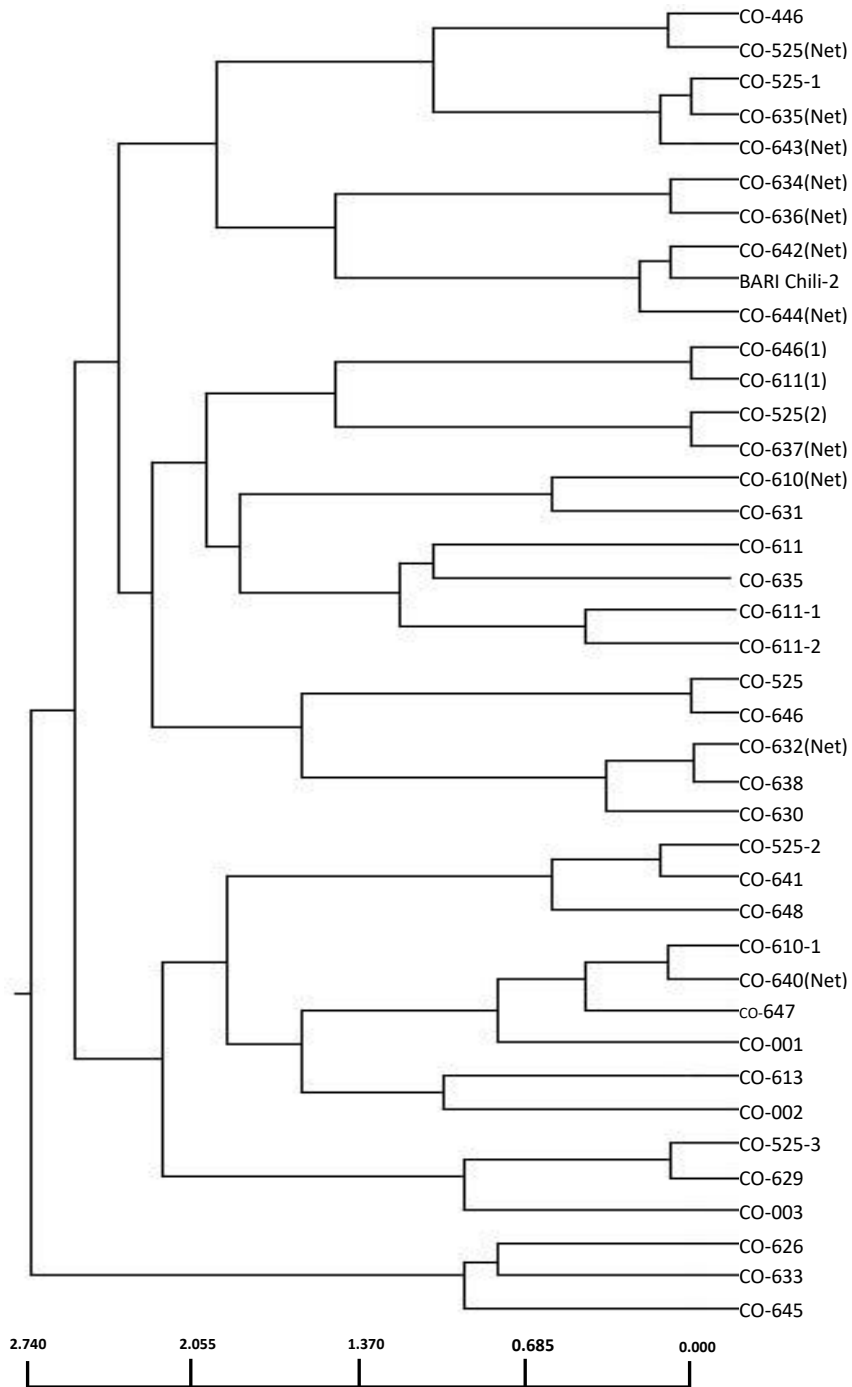


Figure 1. UPGMA dendrogram based on Nei's (1972) genetic distance

Table 6. Summary of heterozygosity statistics for all loci

Locus	Obs. Hom.	Obs. Het.	Exp. Hom.*	Exp. Het.*	Nei**	Ave. Het.
CAMS-075	0.949	0.051	0.289	0.711	0.702	0.025
CAMS-864	1.000	0.000	0.387	0.613	0.605	0.000
CAMS-880	0.917	0.083	0.360	0.641	0.632	0.038
CAMS-065	0.892	0.108	0.357	0.643	0.634	0.050
CAMS-236	1.000	0.000	0.266	0.734	0.725	0.000
CAMS-885	0.925	0.075	0.450	0.550	0.543	0.038
CAMS-647	1.000	0.000	0.256	0.744	0.735	0.000
CAMS-855	1.000	0.000	0.306	0.694	0.685	0.000
Mean	0.960	0.040	0.334	0.666	0.658	0.019
St. Dev.	0.045	0.045	0.067	0.067	0.066	0.021

* Expected homozygosity and heterozygosity were computed using Levene (1949)

** Nei's (1973) expected heterozygosity

CONCLUSION

The study revealed presence of enough genetic variability among the germplasm at molecular level. The first four principal components accounted for 72.44% of the total genetic variance among the accessions. A total of 35 alleles with mean PIC of 0.42 obtained from the molecular analysis show the informative nature of SSR primers and their superiority in genetic diversity assessment. The high amount of genetic variability established by the SSR primers is an indication of the high amount of additive genetic variance within the population. This implies that substantial progress can be made through hybridization. On cluster analysis, the genetic diversity among the cluster one and cluster three was lesser compared to cluster two. It may provide the basic information to the breeders or researchers to select parent in the breeding programme or other programme as per their objectives.

REFERENCES

- Akatas, H., Abak, K. and Sensoy, S. (2009). Genetic diversity in some Turkish pepper (*Capsicum annuum* L.) genotypes revealed by AFLP analyses. *African Journal of Biotechnology*, 8: 4378-4386.
- BBS. (2016). Bangladesh Bureau of Statistics. Statistics and Information Division. Ministry of Planning. Government of the Peoples Republic of Bangladesh. www.bbs.gov.bd.
- Dellaporta, S.L., Woo, J. and Hicks, J.B. (1983). A plant DNA miniprep: version II. *Plant Molecular Biology Reporter*, 1: 19-21.
- DeWoody, J.A., Honeycutt, R.L. and Skow, L.C. (1995). Microsatellite markers in white-tailed deer. *Journal of Heredity*, 86: 317-319.
- Geleta, L.F, Labuschagne, M.T. and Viljoen, C.D. (2005). Genetic variability in pepper (*Capsicum annuum* L.) estimated by morphological data and amplified fragment length polymorphism markers. *Biodiversity and Conservation*, 14: 2361-2375.

- Gupta, P.K., Balyan, H.S., Sharma, P.C. and Ramesh, B. (1996). Microsatellites in plants: A new class of molecular markers. *Current Science*, 70: 45-54.
- Hamrick, J.L. and Godt, M.J.W. (1996). Conservation genetics of endemic plant species. In: Avise JC, Hamrick JL (Eds.), *Conservation Genetics*. Chapman and Hall, New York. Pp. 281-304.
- Hedrick, P.W. (1999). Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, 53:313-318
- Kochieva, E.Z. and Ryzhova, N.N. (2005). Molecular AFLP Analysis of the Genotypes of Pepper (*Capsicum annuum* L.) cultivars. *Russian Journal of Genetics*, 39:1345-1348.
- Liu, K. and Muse, S.V. (2005). Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, 21: 2128-2129.
- Mimura, Y., Inoue, T., Minamiyama, Y. and Kubo, N. (2012). An SSR-based genetic map of pepper (*Capsicum annuum* L.) serves as an anchor for the alignment of major pepper maps. *Breeding Science*, 62: 93-98.
- Nei, M. (1972). Genetic Distance between populations. *The American Naturalist*, 196: 283-292.
- Pickersgill, B. (1997). Genetic resources and breeding of *Capsicum* spp. *Euphytica*, 96: 129-133.
- Slatkin, M. (1987). Gene flow and geographic structure of natural populations. *Science* (Washington, DC), 236: 787-792.
- Se-Jong, O.H., Song, J.Y., Lee, J., Lee, G.A., Ko, H.C., Stoilova, T., Krasteva, L., Kim, Y.G., Rhee, J.H., Gwag, J.G., Ro, N.Y., Hur, O.S. and Lee, M.C. (2012). Evaluation of genetic diversity of red pepper landraces (*Capsicum annuum* L.) from Bulgaria using SSR markers. *Korean Journal of International Agriculture*, 24: 547-556.
- Tam, S.M., Mhiri, C., Vogelaar, A., Kerkveld, M., Pearce, R.S. and Grandbastien, M.A. (2005). Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. *Theoretical and Applied Genetics*, 110: 819-831.