Annals of Plant and Soil Research 24(4):618-629 (2022)
https://doi.org/10.47815/apsr.2022.10218

Genetic diversity analysis using SSR marker in Naga King chilli (*Capsicum chinense* Jacq.) genotypes

MOATOSHI JAMIR, PANKAJ KUMAR SHAH* AND L. DAIHO**

Department of Genetics & Plant BreedingSchool of Agricultural Sciences & Rural Development Nagaland University, Medziphema-797106

Received: April, 2022; Revised accepted, June, 2022

ABSTRACT

The present investigation has been undertaken with the objective to study the genetic variability and diversity among the Naga king chilli genotypes using SSR marker at Medziphema during kharif season of 2018. All total 89 bands were produced by 30 primers, 37 bands being polymorphic and producing average percentage polymorphic bands of 41.66 %. The Primer CAMS-826 showed highest percentage of polymorphic bands (75%), followed by primer CAMS 032, CAMS-117, CAMS- 864 and CAMS-647 which revealed (66.66 %) and primer CAeMS 009, CAMS 163, CAMS 871, HmpsE 047 AND HpmsE 014 exhibited 50% polymorphism which indicates that the primer can be used for genetic diversity study of C. chinense genotypes. Genetic parameters analyzed using Popgene® software indicated that Peren population has higher diversity among the four region populations. Jaccard's similarity coefficient between the genotypes ranged from 0.64 to 0.96, indicating moderate variation between the genotypes. Clustering analysis depicted three clusters where germplasm collected from Peren and Medziphema showed close relatedness and Principal Coordinate Analysis (PCo A) plot obtained also supports cluster analysis data. Germplasm evaluation is of immense important for genetic crop improvement.

Keywords: Genetic diversity, Genetic variability, King Chilli, SSR marker.

INTRODUCTION

The genus Capsicum consists of five cultivated and twenty-five wild species of them *Capsicum chinense* is immensely popular for its unique aroma and high pungency. Most of the chilli species and varieties cultivated in India contain around 1% capsaicin but Naga chilli has around 2-4% capsaicin (Sanatombi and Sharma, 2008). Capsaicin is mainly used as a spice, as food additive, and in pharmacological applications. For better utilization of genetic resource, it is important to know how genetic variation is distributed, and which characteristics of the species and environment influence this distribution. Morphological markers are simplest approach for the assessment of genetic diversity in crop plants. However, level of polymorphism, for morphological traits in elite genotypes is sometimes too limited and inadequate to allow for genotype favoritism. In recent years, molecular markers have proved to be useful in assessing genetic diversity analysis which are more stable and plays an important role in estimating relatedness between cultivars and in crop improvement programme. Among molecular markers, simple-sequence repeats (SSR) have high reproducibility and better use in germplasm characterization, and genetic diversity analysis in cultivated spp. Simple sequence repeat (SSR) as DNA markers have proved to be useful over other commonly used molecular markers because of their co-dominant, polymorphic and readily transferable nature. In the present study 11 landraces were collected from different hotspots in Nagaland to study the genetic variability and diversity analysis using SSR marker in King Chilli of Nagaland. The study will be helpful to identify the landraces with diverse background and their use in trait specific breeding as well as varietal development programme.

MATERIALS AND METHODS

A total of 11 *Capsicum chinense* landraces (mature fruits) were collected from different hotpot areas of Nagaland, India. These germplasms were sown in the nursery and the seedlings were transplanted in randomized block design (RBD) with three replications at experimental farm SASRD, Medziphema during kharif 2018. Based on the data collected the landraces were again planted in RBD with three

** Department of Plant Pathology, SASRD, Nagaland University, Medziphema
*Corresponding author email: pankajshah0112@gmail.com

replications during kharif 2019. For genetic diversity studies fresh leaves from each landrace were collected, lyophilized at −40 °C for 48 h and then stored at −20 °C for further use. DNA extraction was done following the protocol proposed by Bhau*et al.* (2015). The purity of the extracted DNA was checked by 0.8% agarose gel followed by quantitative analysis using Nano Bio Spectrophotometer (Eppendorf, Germany). Genetic diversity study was performed using 30 SSR primers (Bioserve, USA). PCR reaction mixture of 20 µL was made containing 6 µL DNA (30 ng/µL DNA stock), 1.6 µL each forward and reverse primer, 10 µL Hi- Chrome PCR Master Mix and the final volume adjusted with double distilled autoclaved water. PCR programme was set as follows- initial denaturation of 94 °C for 5 min followed by 35 cycle's denaturation at 94 °C for 1 min, 1 min of elongation at 72 °C and final extension of 72 °C for 10 min. Annealing temperature (±2) of the primers was adjusted depending on primers melting temperature (Tm). On the basis of molecular size (base pair) of the amplified product the amplicons were scored as 1 (present) and 0 (absent) respectively and dendrogram was created using (Rohlf, 2000)NTSys software (Ver. 2.02) followed by creation of PCoA plot. From each gel polymorphism information content (PIC) was calculated as- PIC=1 – Σfi2, where f is the frequency of ith allele (Botstein *et al.*, 1980). Using POPGENE software version 1.32 (Yeh *et al.*, 2000) average number of allele (Na) count, effective number of alleles (Ne), Nei's (1973) gene diversity(He), Shannon's (Shannon and Weaver, 1949) diversity index (I), diversity within population (Hs), total gene diversity (Ht) and gene flow (Nm) were analysed. Genetic diversity within and among the populations (AMOVA) was determined using GenAlex software ver. 6.5. (Peakall and Smouse, 2006).

RESULTS AND DISCUSSION

Marker efficiency of SSR primers

In the present investigation, total of 89 bands were produced by 30 primers, 37 bands being polymorphic, producing average percentage polymorphic bands of 41.66 %. Patel *et al.* (2011) also observed similar percentage of polymorphism (50%) in thirteen cultivars of *C. annum* using SSR marker. The number of polymorphic bands revealed per primer ranged from 1 to 4 with an average of 1.23 bands per primer and average PIC value of 0.22 which is close to the results of Sharmin*et al.* (2018) for an average of 2 bands per primer for *Capsicum* species. In the present study primer CAMS-826 showed highest percentage of polymorphic bands (75 per cent), followed by primer CAMS 032, CAMS-117, CAMS-864 and CAMS-647 showed (66.66per cent) and primer CAeMS 009, CAMS 163, CAMS 871, HmpsE 047 AND HpmsE 014 showed 50 per cent polymorphism which indicates that the primer can be used for genetic diversity study of *C. chinense* genotypes. Based on this, SSR primers used in the experiment can be said to reveal better information as the PIC value obtained in the study was found greater than or equal to 0.5 in primer CAMS 163, CAMS-117, CAMS-864, HpmsE 014 and CAMS-826. The average Rp and MI value calculated for SSR primers was 0.47 and 0.13 respectively. Rp specify the potentiality of highly informative primer for genotypes differentiate and MI measures the usefulness of a marker system. Based on this primer CAMS-647 can be considered most informative primer as it possesses maximum Rp value (1.49) followed by primer CAMS-864 (1.11), CAMS 032 (1.09)CAMS-855 (0.94), HpmsE 014 (0.90) AND CAMS-826 (0.90) while primers CAMS-142, CAMS-313, CAMS-378, CAMS-348, CAMS 228, CAeMS 114, CAMS 212, HpmsE 070, CAMS 071 AND CAMS 066 (0.18) to be least informative primer.

Genetic parameters revealed by SSR primers

Genetic structure parameters such as observed number of allele (Na), effective number of allele (Ne), Nei's Gene diversity (He) and Shanon's diversity index were calculated (Table 2). The mean values of '*Na', 'Ne', 'He'* and '*I'* were recorded maximum for genotypes obtained from Peren which is taken as single population. While lower values were obtained for Mon population. Such difference may occur due to inclusion of maximum number genotypes in Peren population than Mon. The magnitude of all the parameters obtained for the four populations were found to be high, indicating higher polymorphism among the populations.

Table 1: Details on primer sequences, polymorphic bands, percentage polymorphism and Polymorphism Information Content (PIC) for SSR primers

Primer Code	Primer Sequence (5/ - 3/)	No. of polymorphic bands	Percentage of Polymorphic bands (PPB)	Resolving Power (RP)	Marker Index (MI)	Polymorphic Information Content (PIC)
CAeMS-009	ACGCACCAACGAATATCTATCTCA GTTTCCGTCCAGATCTACCTTTCCGC	1	50.00	0.45	0.12	0.25
CAeMS-068	ATCAAATCTCAACACATGGTGGCT GTTTACTGTATCTCCGGCCCTGTCA	1	33.33	0.36	0.15	0.47
CAMS-142	GAGCGCTTAAGTGGTCATAGG CTACAACGCCCCAAAACAAT	1	33.33	0.18	0.01	0.05
CAMS-163	TCCATATAGCCCGTGTGTGA GCGTGGGAATACAATGCTAGA	1	50.00	0.54	0.01	0.59
CAMS-313	CAGCCTGCTTGGCTAGAACT TCGTCATGCATGGCTAATCT	1	33.33	0.18	0.01	0.05
CAMS-378	GAATCGACGCGTTTCTAGC TGTGGGGAGAGAGAGGAAGA	1	33.33	0.18	0.05	0.17
CAMS-348	CTGAAGTCGGCTAGATGCCTA TCAAAGCTATGGAGGAAAAGGA	1	33.33	0.18	0.01	0.05
CAMS-451	TGCATTGGTGGGCTAACATA GCTCTTGACACAACCCCAAT	1	33.33	0.36	0.03	0.11
CAMS-032	TGCCACATAGGTTGGCTTTC CAAAGCCAATGCACATAATCA	1	66.67	1.09	0.43	0.33
CAMS-094	TGTAGCTCACATCGTCTCCACT GCATTGCATTTCACTTGCAT	2	33.33	0.36	0.03	0.11
CAMS-228	GAGGGCTAAGCAAAGCAGAA TGCATGTTTCCCTTAGTTTCC	1	33.33	0.18	0.01	0.05
CAeMS-144	ATAACTTTGATTCCTAGTTCGGCG GTTTGAACCCCCAATCATCATATCCTCA	1	33.33	0.18	0.01	0.05
CAMS-212	TTCCCTTTCCCAACATGGTA ACACCCGAAGATGGGTTAGA	1	33.33	0.18	0.01	0.05
CAMS-72	CCCGCGAAATCAAGGTAAT AAAGCTATTGCTACTGGGTTCG	1	33.33	0.36	0.03	0.11
CAMS-117	TTGTGGAGGAAACAAGCAAA CCTCAGCCCAGGAGACATAA	2	66.67	00.36	0.73	0.55
CAMS-871	ACAAAGCATCGGCTGAAAAT GCGACCAAGTACCAACAGGT	1	50.00	0.36	0.08	0.17
CAMS-215	CGTGGGTGGTCTAGGATGAT GCTGGCAAGTCACTCTGGAT	1	33.33	0.66	0.10	0.33
At3G-44600	TCCTTTATACCGACTTGAAGCTATTG AGATTCTATGTTTCTTGAAAGCACAGC	1	33.33	0.66	0.10	0.33
CAMS-396	GTCGGCCGTCATTCACTATT ATAACTTTGATTCCTAGTTCGGCG	1	33.33	0.23	0.03	0.11
CAMS-179	CATGTCATGAAGTTGATAAGACAATG TGTTCCAGTGAAAGGCTTCTT	1	33.33	0.23	0.03	0.11
CAMS-855	AAGTGTCAAGGAAGGGGACA CCTAACCACCCCCAAAAGTT	1	33.33	0.94	0.15	0.47
CAMS-864	CTGTTGTGGAAGAAGACGACA GCTTCTTTTTCAACCTCCTCCT	2	66.67	1.11	0.73	0.55
HmpsE-047	AACCCGTGTTCAATCCCCAAAT TGGCCATACCACCAGCAGTAGA	1	50.00	0.54	0.12	0.25
HmpsE-070	CACTCGTTATATTTTTCTGTCTCG GTGAATATATCCGACCCTGTTT	1	33.33	0.18	0.01	0.05
CAMS-071	AATGGGATCTGCATGAGACA TTCCCTAAAAGATGGTGATTCC	1	33.33	0.18	0.01	0.06
HpmsE-014	CTTTGGAACATTTCTTTGGGGG GCGGACGTAGCAGTAGGTTTGG	2	50.00	0.90	0.31	0.62
CAeMS-060	ATCAAGACAACAACATCATGGGGA GTTTCGCCTATCAACAATGGCAAATACA	1	25.00	0.68	0.02	0.09
CAMS-647	CGGATTCGGTTGAGTCGATA GTGCTTTGGTTCGGTCTTTC	2	66.67	1.49	0.21	0.33
CAMS-826	CTTGATCTCAAGAACCAGCTACAA TGTACATTGAAGACACGGAAGAA	3	75.00	0.90	0.50	0.68
CAMS-066	AAAAACATGCACCAGTCCTT CAACCGCCTGAATTTTCTCT	1	33.33	0.18	0.01	0.06
Total	-	37	-	-	-	-
Average		1.23	41.66	0.47	0.13	0.22

Further it was observed that within population the diversity was high compared to the average diversity (1.18). Moses *et al.* (2014) in their study reported high value of gene diversity (*He*=0.58) among the accessions of *C. chinense*using 9 microsatellite markers. Heterozygosity (*Ht*) is the frequency of genes in each locus and higher the heterozygosity greater will be the total diversity. An average heterozygosity value of 0.15 was recorded indicating moderate variation for the studied population. Similar results for heterozygosity were produced by Aguilar *et al.* (2016) in Mexican chili population using microsatellite markers. A '*Gst*' value of 0.45 was observed which is low and on the other hand, a gene flow level of 0.60 was observed which is slightly lower than the threshold value proposed by Slatkin (1987). The present study is assisted by the fact that gene flow level was not very high, resulting in low value of *Gst*, which ultimately decreases the heterozygosity level of individuals in a population. According to Hazarika and Neog (2014) high level of gene flow between *C. chinense* populations might occur during cross pollination or wide dispersal of pollens and seeds by birds and insects.

Table 2: Genetic diversity parameters based on SSR marker in *Capsicum chinense* species under study

C. chinense population ID	Mean *Na*	Mean *Ne*	Mean *He*	Mean *I*
Population 1 (Mon)	1.09	1.06	0.03	0.05
Population 2 (Peren)	1.42	1.36	0.19	0.26
Population 3 (Dimapur)	1.13	1.09	0.05	0.08
Population 4 (Medziphema)	1.10	1.07	0.04	0.06
Average	1.18	1.14	0.07	0.11
Standard deviation (SD)	0.14	0.12	0.07	0.08
Ht	*Hs*	*Gst*	*Nm*	
0.15	0.08	0.45	0.60	

Pop population, Na observed number of alleles, Ne effective number of alleles, He Nei's gene diversity, I Shanon's diversity index, HT total species diversity, Hs diversity within population, Gst coefficient of gene differentiation, Nm gene flow.

Table 3: Analysis of molecular variance (AMOVA) based on SSR marker in *C. chinense*Jacq.

Source of variation	Degrees of freedom (df)	Sum of squares (SS)	Variance component	Percentage of total variance	P value
Among pops	3	3.92	0.07	6	0.058
Within pops	7	7.90	1.13	94	0.039
Total	10	11.82	1.20	100	

P value = probability of obtaining a more extreme component by chance alone, Pop=population

Analysis of molecular variance (AMOVA) of *C. chinense*

AMOVA computed for the four populations showed significant variance within populations, which is about 94%, while not much variation was observed among the populations (6%). The variance within the population was found to be 1.13, which is high than the variance among the populations (0.07). This revealed that majority of the variation occurred mainly within the population due to higher level of gene flow between populations rather than the variation caused by geographical distance. Similar results have been reported by Baruah *et al.* (2017).

Genetic relationship among different populations of *C. chinense*

Jaccard's similarity coefficient ranged from 0.60 to 0.96 with a mean value of 0.78, indicating sufficient genetic variations among the studied populations. Sharmin*et al.* (2018) observed a similarity range of 0.000-1.000 in *Capsicum* sp. with 0.5 mean value suggesting prevalence of average genetic variation in the

studied material. Cluster analysis revealed formation of three main clusters for four population of *C. chinense* Jacq. where, cluster 1 includes four genotypes collected from Peren and Dimapur respectively. Cluster 2 showed five genotypes collected from Dimapur, Mon and Peren. Meng *et al.* (2017) also observed similar results. Hossain *et al.* (2014) also observed formation of three major clusters in diversity study of 22 chilli germplasm. It can be seen that for majority of the genotypes, the clusters were formed according to their place of collection, suggesting the presence of sufficient genetic variation for the populations. Further it can be observed that three genotypes collected from Dimapur, one genotype collected from Dimapur and one genotype collected from Mon formed cluster with Peren population. Such condition may happen due to duplicity in germplasm occurred during collection of fruits from vendor or the primers used in the study are not sufficient to differentiate these genotypes. Hanacek*et al.* (2009) also reported the possibility of germplasm duplication as the genotypes which are thought to be distantly related showed similar characteristics.

Table 4: Nei's genetic identity (above diagonal) and genetic distance (below diagonal) among four populations of *C. chinense*

Pop. ID	POP 1 (Mon)	POP 2 (Peren)	POP 3 (Dimapur)	POP 4 (Medziphema)
POP 1 (Mon)	****	0.8981	0.9335	0.8940
POP 2 (Peren)	0.1074	****	0.8950	0.8724
POP 3 (Dimapur)	0.0688	0.1110	****	0.9174
POP 4 (Medziphema)	0.1121	0.1365	0.0862	****

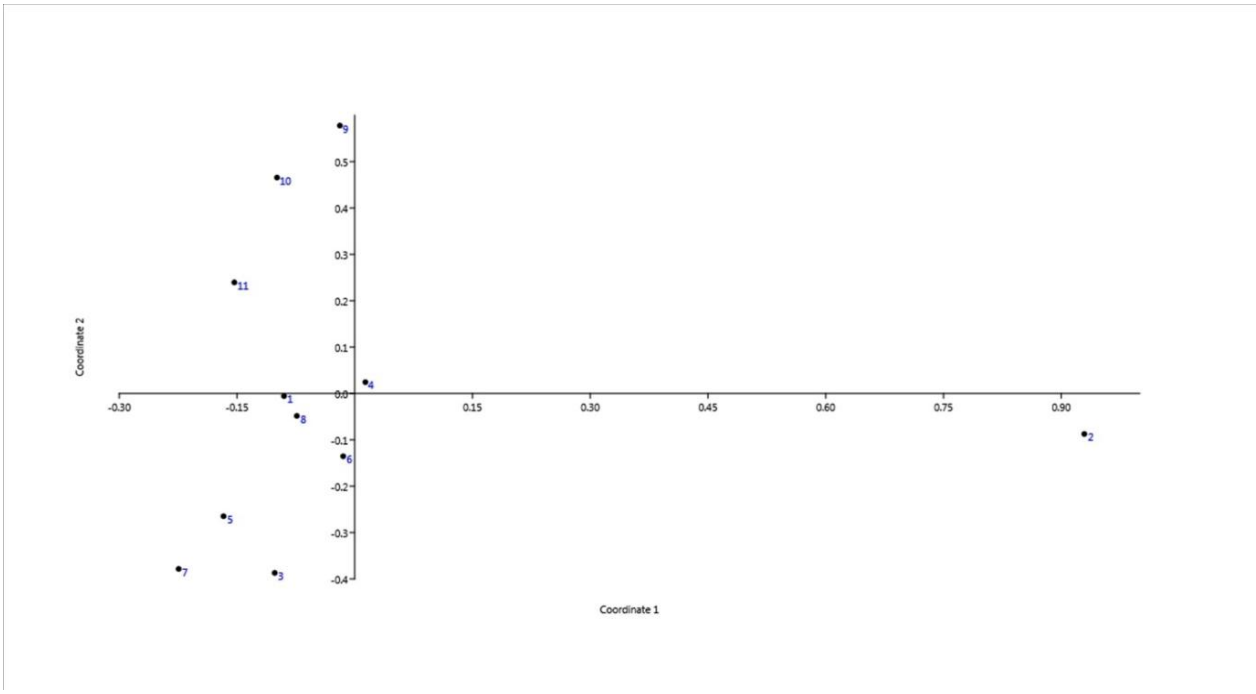

Fig. 1: 2-D PCoA analysis of *Capsicum chinense* germplasm based on SSR marker data

Among populations, Nei's genetic identity ranged from 0.872- 0.933, with Pop 3 and 1 showing highest similarity and Pop 4 and 2 showing least similarity. Between the populations the genetic distance value ranged from 0.068- 0.136, with Pop 2 and 4 shows maximum and Pop 1 and 3 minimum genetic distances. According to Zhang *et al.* (2016) such collinearity may exist due to large sample size of the studied material which is sufficient to provide its complete genetic relation. The result obtained in the present study could be used as a baseline data for selection of diverse parents for further traits specific breeding programme.

Principal coordinates analysis (PCoA) was carried out to further validate the data obtained from cluster analysis (Fig. 1). The two-dimensional principal coordinate's analysis drawn compliments cluster analysis with clear formation of three separate groups. The genotypes collected from Dimapur and Mon that showed grouping with Peren population in cluster analysis also showed separation from their respective group in 2-D PCoAplot. This can be attributed to the fact that the conclusion drawn from cluster analysis is fully supported by the results of PCoA plot. Meng *et al.* (2017) also reported similar conclusions for both cluster analysis and principal coordinate's analysis.

The present study showed that SSR analysis is quick and reliable procedure for diversity analysis with sufficient polymorphism within the experimental population. From the present investigation high genetic variability has been observed within the four populations and among them Peren and Medziphema population have the farthest distance. The results would be certainly helpful for the breeders for genetic enhancement of the species.

REFERENCES

Aguilar, T. R.,Higinio, L. S.Amalio, S., Ernestina, V. M., Pedro, L., Víctor, A.R., Víctor, G. H.and Humberto, V. H. (2016) Characterization of geneticdiversity of native 'Ancho' chili populations of Mexicousing microsate llite markers. *Chilean journal of agricultural research* **76**:18-26.

Baruah, J., Gogoi, B., Das, K., Ahmed, N.M., Sarmah, D.K., Lal, M. and Bhau, B.S., (2017)Genetic diversity study amongst Cymbopogon species from NE-India using RAPD and ISSR markers. *Industrial Crop Production* 95:235–243.

Bhau, B.S., Gogoi, G., Baruah, D., Ahmed, R., Hazarika, G., Borah, B., Gogoi, G., Sharmah, D.K., Nath, S.C., Wann, S.B., (2015) Development of an effective and efficient DNA isolation method for *Cinnamomum* species. *Food Chemistry*. 188: 264–270.

Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics***32**(3): 314–331.

Hanacek, P. Vyhnanek, T., Rohrer, M., Sprojcarova, Jaroslava and Stavelikova, H. (2009) DNA polymorphism in genetic resources of red pepper using microsatellite markers. *Horticultural Science* 36: 127-132.

Hazarika, R., Neog, B., (2014) Evaluation of genetic diversity in Bhut Jolokia (*Capsicum chinense*Jacq.) accessions using ISSR marker. *International Journal of Basics Applied Science Research***1**(1):72–79.

Hossain, S. M., Habiba, U., Bhuyan, S. I., Haque, M. S., Begum, S. N. and Hossain, M. D. (2014) DNA fingerprinting and genetic diversity analysis of chilli germplasm using microsatellite markers. *Biotechnology* 13: 174-180.

Meng, C., Wei, X., Zhao, Y., Yuan, Y., Yang, S., Wang, Z., Zhang, X., Jiuwen, S., Zheng, X., Yao, Q. and Zhang, Q. (2017). Genetic diversity analysis of capsicum genus by SSR markers. *Molecular Plant Breeding*. 8: 70-78.

Moses, M., Umaharan, P., and Selvadurai, D. (2014). Microsatellite based analysis of the genetic structure and diversity of *Capsicum chinense* in the Neotropics. *Genetic Resources and Crop Evolution*. 61 (4): 741-755.

Nei, M., (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*. 70: 3321–3323.

Patel, A. S., Sasidharan, N., Ashish, G. V. and Kumar, V. (2011) Genetic relation in *Capsicum annum* [L.] cultivars through micro satellite markers: SSR and ISSR. *Electronic Journal of Plant Breeding***2**(1): 67-76.

Peakall, R. andSmouse, P.E. (2006)GenAlEx: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–295.

Rohlf, F.J., (2000) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System Version 2.1. Exeter Software, Setauket, NY.

Sanatombi, Keithellakpam& Sharma, Gurumayum. (2008) Capsaicin Content and Pungency of Different Capsicum spp. Cultivars. *NotulaeBotanicae Horti Agrobotanici Cluj-Napoca.* 36 (2): 89-90.

Shannon, C.E. and Weaver, W. (1949) The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Sharmin, A., Hoque, E.M., Haque, M.M., Khatun, F., (2018) Molecular diversity analysis of some chilli (Capsicum spp.) genotypes using SSR markers. *American Journal of Plant Sciences.* 9: 368–379.

Slatkin, M. (1987) Gene flow and geographic structure of natural populations. *Science* 236:787–792.

Yeh, F.C., Yang, R.C. and Boyle, T., (2000) POPGENE 1.32: A Free Program for the Analysis of Genetic Variation Among and Within Populations Using Co-dominant and Dominant Markers. Dept of Ren. Res. Uni Alberta, Canada.

Zhang, X., Zhang, Z., GU, X., Mao, S., Li, X., Chadœuf, J., Palloix, A., Wang, L. and Zhang, B. (2016) Genetic diversity of pepper (Capsicum spp.) germplasm resources in China reflects selection for cultivar types and spatial distribution. *Journal of Integrative Agriculture*. 2016. **15** (9): 1991–2001.