Advances in Bioresearch
Adv. Biores., Vol 11 (2) March 2020: 68-74
©2020 Society of Education, India
Print ISSN 0976-4585; Online ISSN 2277-1573
Journal's URL:http://www.soeagra.com/abr.html
CODEN: ABRDC3
DOI: 10.15515/abr.0976-4585.11.2.6874

Advances in Bioresearch

ORIGINAL ARTICLE

Genetic diversity analysis among important rice (*Oryza sativa* L.) genotypes using SSR markers

Rajesh Yogi[1*], Naveen Kumar, [1] Ravinder Kumar[2] and R.K. Jain[1]
1. Department of Molecular Biology, Biotechnology and Bioinformatics,
CCS Haryana Agricultural University, Hisar - 125004, (Haryana), India
2. Department of Bio and Nano Tecchnology, GJUST, Hisar, Haryana
*Corresponding author: Dr. Rajesh Yogi, Email: rajeshyogi999@gmail.com

ABSTRACT
The study of genetic diversity among rice genotypes is necessary for varietal identification, better understanding of rice evolution and conservation of germplasm. Molecular markers provide novel tools for diversity analysis among various rice genotypes. A DNA fingerprint database of thirteen rice genotypes was prepared using 129 SSR markers. A total of 355 alleles were identified in thirteen rice genotypes. STRUCTURE analysis was performed to estimate the number of populations and two populations were assumed. The allelic diversity was used to construct a dendrogram and two major clusters were observed with the similarity coefficient of 0.53. Analysis of molecular variance was performed to access the genetic variation among groups and among genotypes and found 96.428% variation among genotypes and 3.57% among groups. Two dimensional PCA scaling also exhibited that these genotypes were interspersed into two groups. The observed results indicated good fit genetic diversity in the rice genotypes.
Keywords: *Basmati rice, genetic diversity, SSR*

Received 16.01.2020 Revised 20.02.2020 Accepted 08.03.2020

How to cite this article:
R Yogi, N Kumar, R Kumar and R.K. Jain. Genetic diversity analysis among important rice (*Oryza sativa* L.) genotypes using SSR markers. Adv. Biores., Vol 11 (2) March 2020: 68-74

INTRODUCTION

Rice (*Oryza sativa* L.) stands on the foremost position amongst the food crops and more than 3.5 billion people worldwide depend on it [27]. It provides 43% of the calorie requisite for more than 50% of Indian population. India has 44 million hectares under rice cultivation and ranks second, after China with annual production of about 104 million tons [3]. *Oryza sativa* comprises two major subspecies: the non-sticky, long to short, slender and somewhat flat grained variety *indica* (indica rice) and the sticky, short and roundish grained variety *japonica* (japonica rice); and. Both vary in terms of adaptation related to different geographic, cultural and climatic conditions. The genetic status of Basmati rice varieties regarding *indica* and *japonica* in *Oryza sativa* L. is still not clear.

Primary data in form of genetic diversity is needed for the crop improvement because it helps in the selection of parents with good fit genetic diversity with unique characters and to develop improved recombinants [18, 16]. So, to get improved rice cultivars there is need to study the genetic diversity. It is one of the most diversified crop species as a main cereal crop, because it can adapt in the wide range of ecological, climatic and geographical regions. Glaszmann [9] described clustered or grouped basmati rice cultivars in another group than aerobic *indica* rice and lowland varieties.

Simple Sequence Repeat (SSR) markers are one of the important and effective genomic tools for the assessment of genetic variability and have been efficiently employed for various purposes which includes genome mapping [28, 16], purity testing and identification of varieties [18] and also to resolve of the genetic association amongst different sub-species [19].

The current study was conducted to assess the genetic variability and relatedness amongst thirteen rice genotypes including lowland Basmati (Pusa 1121, Pusa 1460, Basmati 370, Pusa Basmati 1, HBC19 and Improved Basmati 370) and *indica* (PAU201, HKR47 and IR64) and aerobic (MAS25, MAS-ARB25, MAS26 and MAS109)generated by SSR markers. Molecular marker study will help in the classification and

discrimination of genotypes with diverse genetic framework. The observed results will help in the selection of dissimilar parents for crossing programs and also helps in development of the germplasm base for various breeding programs.

MATERIAL AND METHODS

Plant material

A total of thirteen rice genotypes used in the present study were grouped into three groups *viz*. group I-aerobic (MAS25, MASARB25, MAS26 and MAS109) group II- lowland Basmati (Pusa 1121, Pusa 1460, Pusa Basmati 1, Basmati 370, HBC19 and Improved Basmati 370) and group III-*indica* (PAU201, HKR47 and IR64). The rice genotypes were grown in the pot-house at Department of Molecular Biology, Biotechnology, and Bioinformatics, CCS Haryana Agricultural University, Hisar.

SSR analysis

Leaf samples were collected from 2-3-week old seedlings of the genotypes grown in pot-house. Genomic DNA was isolated from thirteen rice genotypes using CTAB method (Saghai-Maroof et al. 1984) and dissolved in TE buffer (0.1 mM EDTA,10 mM Tris, pH 8.0). The isolated DNA was checked for its quantity and quality by using 1% (w/v) agarose gel electrophoresis. PCR amplifications were performed in a reaction volume of 20µl containing primer 50-75ng template DNA,100 µM dNTPs,1.0X PCR buffer,1.5 units *Taq* DNA polymerase and 0.3µM of each. For amplification, PCR amplification was carried out using the thermal profile: one cycle of initial denaturation at 94°C for 5 minutes, thirty five cycles for initial denaturation at 94°C for 1 minute, primer annealing at 55°C for 1 minute and extension at 72ºC for 2 minutes and one cycle of final extension at 72ºC for 7 minutes followed by storage at 4ºC for infinite time. The amplified products of SSR markers were resolved on 4% polyacrylamide gels with ethidium bromide staining, and good quality gels were photographed under ultra violet light region. On the basis of electrophoretic mobility relative to 20bp DNA ladder (*i.e.* known molecular weight markers), the size of the amplified bands for each microsatellite markers was analyzed.

Allele scoring and data analysis

The visible bands of each genotype were scored as binary data, the marker alleles were coded '0' for the absence and '1' for the presence of the allele and entered in the excel spreadsheets. The alleles with negligible intensity, diffused or smeared data was excluded from the final data analysis and considered as missing data [15]. The SSR bands amplified by using given SSR primers were further treated as a unit character. The frequency of microsatellite marker polymorphism was considered on the presence or absence of shared bands. The basic statistics including the numbers of amplicons, polymorphic information content (PIC) and major allele frequency were calculated for each primer to analyze the efficacy of SSR markers. Polymorphic information content (PIC) for each marker was calculated according to Anderson *et al.* [2]. The polymorphic information content (PIC) values of individual primer were calculated based on the formula as given:

$PIC = 1- \sum^{n}i = 1\ P^{2}ij$ [16]

Phylogenetic trees from the genotyping data of SSR markers were constructed. The Dice coefficient was calculated based on dissimilarity matrix and the cluster analysis was performed based on the unweighted neighbor-joining method. To get the better reliability of the construction 1000 bootstrapping were used with DAR win (Version 6.0.14) software program (http://darwin.cirad.fr/darwin/Home.php) [17]. For the evaluation of the genetic variation among the population and within the population, the non-parametric analysis of AMOVA was done *via* GenAlEx (version 6.5) [21]. The two dimensional PCA was also constructed for accurately testing the relationships among genotypes using the EIGEN program analysis using AMOVA *via* GenAlEx.

Bayesian clustering methods are very useful computational tools meant for assessment of several features of population. STRUCTURE (version 2.3.4), that is a Bayesian clustering software which identify subpopulations, assign individual to different populations and hybrid zones on the basis of allele frequencies of genotypes [22]. For the estimation of the most likely K, all possible K's (*i.e.,* K = 1 to 10) were evaluated using simulation of 10 iterations, with each iteration consisted of 10,000 'burnin' followed by 100,000 Markov Chain Monte Carlo (MCMC) replications, with default settings for both the Ancestry Model (Admixture Model) and the Frequency Model. The data generated by the STRUCTURE were harvested by using Evanno's method [7]) based program STRUCTURE HARVESTER developed by Earl and Vonholdt [6]. The *ad-hoc* statistical method was used to determine the number of sub-populations (ΔK), based on the rate change in the log probability of data between successive K values.

RESULTS

Microsatellite marker analysis

A total of 129 polymorphic SSR markers were used for diversity analysis of lowland Basmati (Pusa 1121, Pusa 1460, Pusa Basmati 1, Basmati 370, HBC19 and Improved Basmati 370), *indica* (PAU201, HKR47 and IR64) and aerobic (MAS25, MAS-ARB25, MAS26 and MAS109) rice genotypes, which are generally distributed on 12 rice chromosomes. A DNA fingerprint database of thirteen rice genotypes was prepared using 129 polymorphic microsatellite markers. Representative pictures with polymorphism were shown in Fig. 1. A total of 355 alleles with an average number of 2.75 alleles per locus were identified in thirteen rice genotypes. Eight (RM186, RM233A, RM258, RM262, RM269, RM467, RM475 and RM6925) of 129 SSR markers amplified five alleles which were maximum in thirteen rice genotypes. Polymorphic information content (PIC) values of the 129 microsatellite markers based on the allelic patterns of all the genotypes were accessed to determine the efficacy of the primers. The PIC value of SSR markers ranged from 0.124-0.769 with a mean value of 0.479, 53 polymorphic SSR markers have PIC value (> 0.500) with an average of 0.606.

Genetic characterization of thirteen rice genotypes

The unweighted neighbour-joining tree was constructed using dissimilarity data matrix based on Dice coefficient. The unweighted neighbour-joining method with 1000 bootstrapping was adopted to get better results (Fig 2). All the thirteen rice genotypes clustered in two major groups with the similarity coefficient of 0.53. Major group I consisted of aerobic and *indica* rice genotypes while the major group II had Basmati rice genotypes with one aerobic rice variety. The *indica* variety PAU 201 remained ungrouped, which indicated this variety is different from all the rice variety used in the present study. The clustering pattern of the dendrogram followed the initial grouping of the genotypes (aerobic, *indica* and Basmati rice) with little interference.

Analysis of molecular variance was performed for SSR binary data to estimate the genetic variability among genotypes and among groups. AMOVA analysis indicated that the percentage of variation observed 3.57% among groups and 96.428% among the varieties (Table 1).The two dimensional PCA was also constructed for accurately testing the relationships among thirteen rice genotypes using the EIGEN program analysis using AMOVA *via* GenAlEx. The PCA was carried out using the collected binary data of SSR markers (Fig 3). The populations were coded as follows: the aerobic rice genotypes in blue color, Basmati genotypes in red color and the *indica* genotypes in green color. The PCA analysis showed the similar distribution pattern like the pattern observed from the dendrogram. The observed results indicated a good fit genetic diversity among the collected rice varieties.

The population structure analysis was carried out using STRUCTURE software which is Bayesian clustering software that distributes the individuals to the possible population groups on the basis of the allele frequencies of genotypes. Without the prior knowledge of the number of populations, STRUCTURE distributed the thirteen rice varieties in two clusters. The value of ΔK was observed maximum when K=2 (Fig. 4a). Two populations were assumed by the software on the basis of alleles and also supported the initial grouping of the genotypes based on different varieties with little interference (Fig. 4b).

DISCUSSION

A total of 129 polymorphic microsatellite markers widely distributed on 12 rice chromosomes were used to assess genetic diversity among thirteen rice genotypes. As many as 355 alleles were identified in thirteen rice genotypes with an average number of 2.75 alleles per locus. These results were in agreement with Cho *et al*. [5], they observed an average of 2.0-5.5 alleles per locus in used rice germplasm. Eight microsatellite markers amplified maximum five numbers of alleles among thirteen rice genotypes. Panaud *et al*. [20] also reported 2 to 9 alleles for microsatellite markers in 22 *indica* and *japonica* cultivars. High level of polymorphism was observed among thirteen Basmati, *indica* and aerobic rice varieties. Data showed that there was huge diversity among the aerobic and high-yielding *indica*/Basmati rice genotypes. Fifty three out of 129 polymorphic SSR markers used had PIC values >0.50.Markers were classified as informative when PIC was ≥ 0.5.

The unweighted neighbor-joining tree was constructed using dissimilarity data matrix based on Dice coefficient. All the thirteen rice genotypes clustered in two major groups. Major group, I consisted of aerobic and *indica* rice genotypes while the major group II had Basmati rice genotypes. Genetic relationships among these rice genotypes were also assessed by PCA analysis which exhibited that thirteen rice genotypes were interspersed in the two groups. Phylogenetic analysis using microsatellite DNA markers leads to the grouping of most of the Basmati rice varieties into a separate group distinct from those of *indica* and *japonica* rice varieties [12, 18, 14].

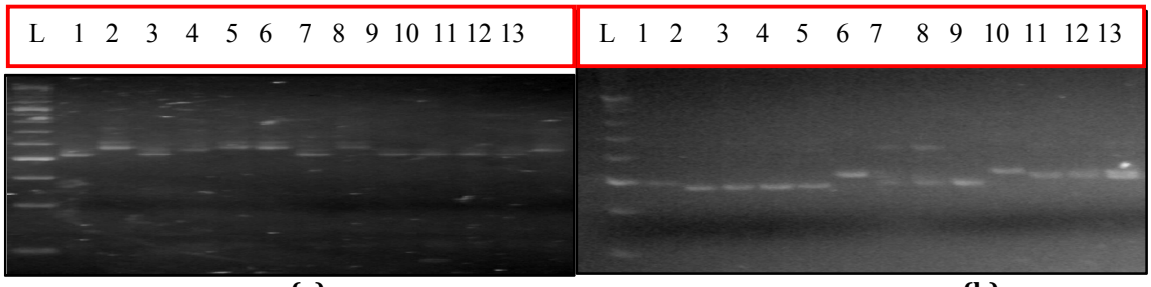
The mean similarity coefficient of aerobic with *indica* (0.652) was higher than those between aerobic and lowland Basmati (0.584) rice and between *indica* and Basmati (0.549) rice genotypes. On the basis of average similarity coefficient values of 129 SSR markers, the aerobic group was found closer to lowland *indica*(Group I, Glaszmann, [9]) than Basmati (Group V, Glaszmann, [9]) rice varieties. Girish *et al*. [8] placed aerobic rice genotypes closer to *indica*than the Basmati rice group. The diversity observed between lowland *indica* (Group I Glaszmann, [9]) and aerobic *indica* (Group II was also known as 'aus' varieties) is in agreement with Glaszmann [9] classification system of different rice genotypes. In most of studies, Basmati types clustered into a separate group distinct from that of *indica* and *japonica* rice varieties [1, 12].

Glaszmann [9] on the basis of allelic variation in fifteen isozyme loci divided Asian rice cultivars into six groups, with two major groups (group I and VI), two minor groups (groups II and V) and two satellite groups (groups III and IV). Group I corresponds to *indica* and group VI corresponds to *japonica*. Group II, III. IV and V were atypical but also classified as *indicas* in the conventional classification. Group V includes aromatic rice of Indian subcontinent. Upland rice varieties (also known as 'aus' varieties) which are early maturing and drought tolerant, have been classified in varietal group II Glaszmann [9]. In this study, clustering or grouping of Basmati rice varieties in a different group than lowland and aerobic *indica*rice is in the conformity with the earlier reports [9, 12]. It must be noted that aerobic rice varieties were developed from different *indica*× upland rice crosses and may have a differential level of genetic content from the recurrent *indica* and donor upland rice varieties.

In a similar study Jasim *et al.* [13] studied genetic diversity in fifty accessions of aromatic rice and three control varieties. The accessions were collected from three different regions. Thirty two polymorphic SSR markers were used and generated 2 to 7 alleles per locus with an average of 4.09 alleles per primer which is higher than the present study. PIC value ranged from 0.25 to 0.98. AMOVA analysis showed the percentage of variation 89% among the population and 11% among the groups in agreement with the results observed in present study. Similarly, Salem *et al.* [26] employed twenty two lines of rice from three regions *i.e.* India (two), Philippines (six), and Egypt (fourteen) for the assessment of genetic diversity with twenty three SSR markers. Population structure analysis revealed that the value of ΔK was found maximum when K=2. Two populations were assumed by the software and clustered the genotypes in two clusters with interference. No clear demarcation was there according to the geographical region and all parameters were found similar to the results observed in present study. Observed results indicated the significant genetic diversity in the rice genotypes. The differences in the results may be due to the use of different primer set and different germplasm.

Table 1. Analysis of molecular variance (AMOVA) obtained in thirteen rice genotypes using SSR markers

Source of variation	Degree of freedom	Variance component	Percentage of variation (%)	P-value
Among groups	2	2.472	3.57	< 0.001
Among genotypes	10	66.750	96.428	< 0.001



(a) **(b)**

Figure 1. Polyacrylamide gel showing allelic polymorphism among 13 rice varieties at (a) RM44 and (b) RM 17524 locus. (Lane L-20 bp ladder; 1-13 rice genotypes --MAS25, MASARB25, IB370, PAU201, MAS26, PUSA1121, MAS109, PUSA1460, PB1, Basmati 370, HKR47, HBC19 and IR64, respectively)

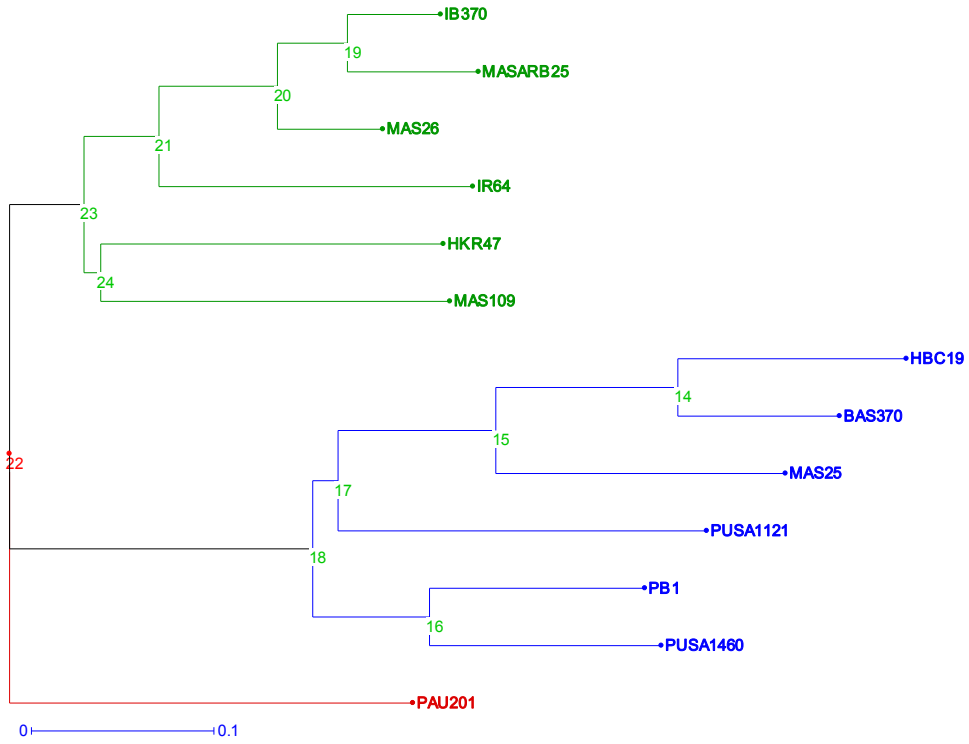


Figure 2. Unweighted neighbor joining display diversity among 13 rice genotypes

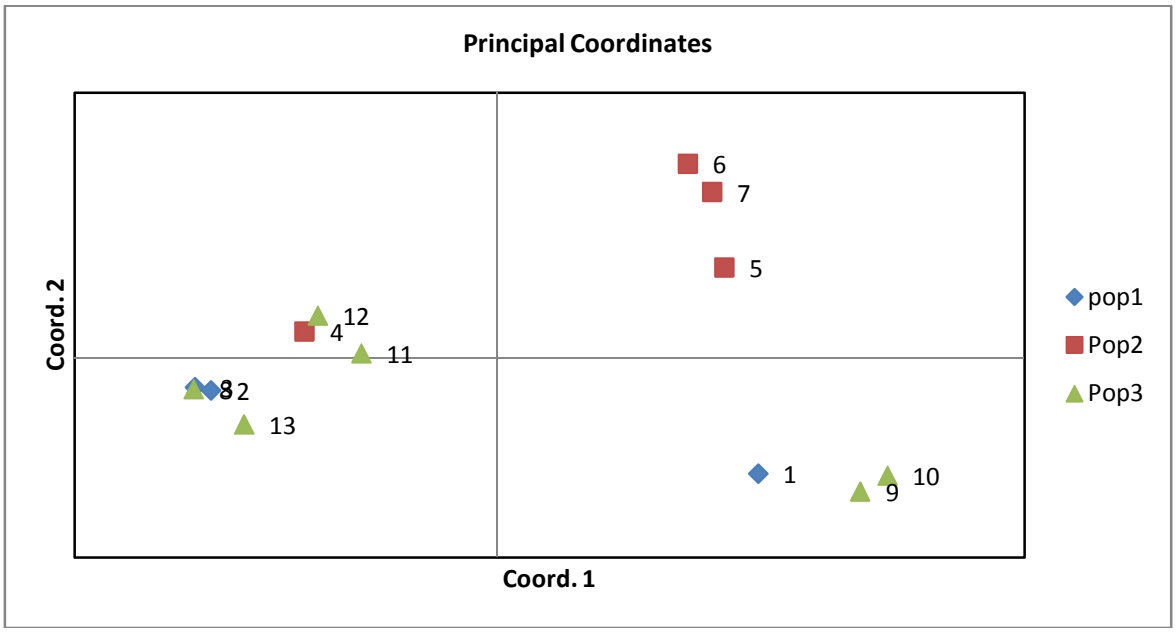


Figure 3. Two dimensional PCA scaling display diversity among 13 rice genotypes, using allelic diversity data at 129 SSR loci

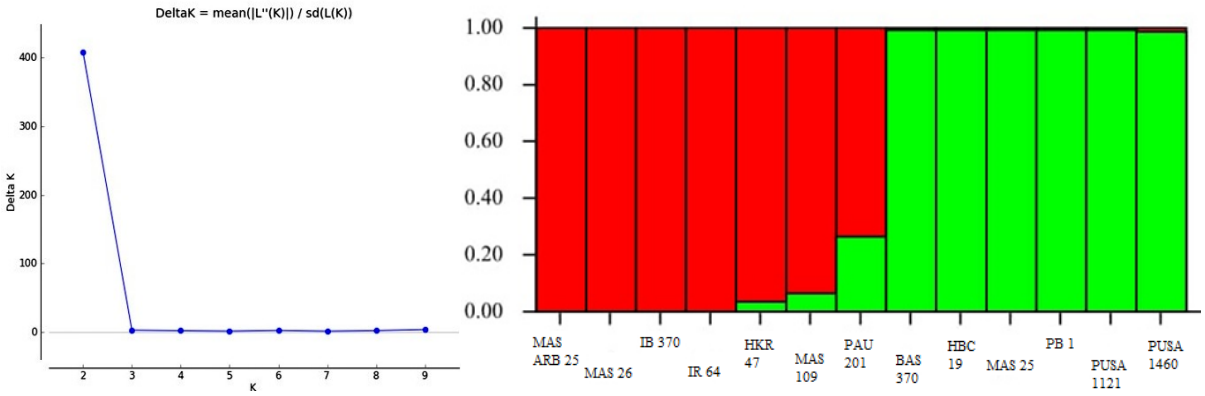


Fig 4a Fig 4b

Figure 4. (a) Relationship between K and ΔK based on STRUCTURE analysis. (b) STRUCTURE analysis of rice genotypes based on SSR data showing a grouping of genotypes when K=2

REFERENCES

1. Aggarwal, R.K., Shenoy, V.V., Ramadevi, J., Rajkumar, R., Singh, L. (2002). Molecular characterization of some Indian Basmati and other elite rice genotypes using fluorescent-AFLP markers. Theoretical and Applied Genetics 105, 680-690.
2. Anderson, J.A., Churchill, G.A., Autique, J.E., Tanksley, S.D., Sorrells, M.E. (1993). Optimizing parental selection for genetic linkage maps. Genome 36, 181-186.
3. Babu, B.K., Meena, V., Agarwal, V., Agrawal, P.K. (2014). Population structure and genetic diversity analysis of Indian and exotic rice (*Oryza sativa* L.) accessions using SSR markers. Molecular Biology Reporter41(7), 4328-4339.
4. Bligh, H.F.J., Blackhall, N.W., Edwards, K.J., McClung, A.M. (1999). Using amplified fragment length polymorphisms and simple sequence length polymorphisms to identify cultivars of brown and white milled rice. Crop Science39, 1715-1721.
5. Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., Park, W.D., Ayres, N., Cartinhour, S. (2000). Diversity of microsatellites derived from genomic libraries and gene bank sequences in rice (*Oryza sativa* L.). Theoretical and Applied Genetics 100, 713-722.
6. Earl, D.A. and VonHoldt, B.M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4(2), 359-61.
7. Evanno, G., Regnaut, S., Goudet, J. (2005). Detecting the number of clusters of individuals using the software Structure: A simulation study. Molecular Ecology. 14(8), 2611-20.
8. Girish, T.N., Gireesha, T.M., Vaishali, M.G., Hanamareddy, B.G., Hittalmani, S. (2006). Response of a new IR50/ Moroberekan recombinant inbred population of rice (*Oryza sativa* L.) from an *indica* × *japonica* cross for growth and yield traits under aerobic conditions. Euphytica. 152, 149-161.
9. Glaszmann, J.C. (1987). Isozyme and classification of Asian rice varieties. Theoretical and Applied Genetics 74, 21-30.
10. Herrera, T.G., Duque, D.P., Almeida, S.P., Nunez, G.T., Pieter, A.J., Martinez, C.P., Tohme J.M. (2008). Assessment of genetic diversity in Venezuelan rice cultivars using simple sequence repeats markers. Electronic Journal of Biotechnology 11(5), 1-14.
11. Jain, N., Jain, S., Saini, N., Jain, R.K. (2006). SSR analysis of chromosome 8 regions associated with aroma and cooked kernel elongation in Basmati rice. Euphytica. 152, 259-273.
12. Jain, S., Jain, R.K., McCouch, S.R. (2004). Genetic analysis of Indian aromatic and quality rice germplasm using panels of fluorescently labelled microsatellite markers. Theoretical and Applied Genetics 109, 965-977.
13. Jasim, A.S., Rafii, M.Y., Latif, M.A., Sakimin, S.Z., Arolu, I.W., Miah, G. (2018). Genetic diversity of aromatic rice germplasm revealed by SSR markers. BioMed Research International. 2018.
14. Kovach, M.J., Calingacion, M.N., Fitzgerald, M.A., McCouch, S.R. (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). Proceedings of the National Academy of Sciences of the United States of America. 106, 14444-14449.
15. Kumar, R., Naik, P.K., Kumar, A., Aggarwal, H., Kumar, A., Chhokar, V. (2016). A combined approach using RAPD, ISSR and bioactive compound for the assessment of genetic diversity in *Aloe vera* (L.) Burm. f. Indian Journal of Biotechnology. 15, 538-49.
16. McCouch, S.R., Teytelman, L., Xu, Y.B., Lobos, K.B., Clare, K., Walton, M., Fu, B.Y., Maghirang, R., Li, Z.K., Xing, Y.Z., Zhang, Q.F., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D., Stein, L. (2002). Development and mapping of 2,240 new SSR markers for rice (*Oryza sativa* L.). DNA Research6, 199-207.
17. Moges, A.D., Admassu, B., Belew, D., Yesuf, M., Njuguna, J., Kyalo, M., Ghimire, SR. (2016) Development of microsatellite markers and analysis of genetic diversity and population structure of *Colletotrichum gloeosporioides* from Ethiopia. PloS one. 11(3), e0151257.

18. Nagaraju, J., Kathirvel, M., Ramesh, K.R., Siddiq, E.A., Hasnain, S.E. (2002). Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence based ISSR-PCR and SSR markers. Proceedings of the National Academy of Sciences of the United States of America. 99(1), 5836-5841.
19. Ni, J., Colowit, P.M., Mackill, D. (2002). Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Science 42, 601-607.
20. Panaud, O., Chen, X., McCouch, S.R. (1996). Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice. Molecular Genetics and Genomics. 252(5): 597-607.
21. Peakall, R.O. and Smouse, P.E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Resour. 6(1), 288-95.
22. Pritchard J.K., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics. 155(2), 945-59.
23. Ravi, M., Geethanjali, S., Sameeyafarheen, F., Maheswaran, M. (2003). Molecular marker based genetic diversity analysis in rice (*Oryza sativa* L.) using RAPD and SSR markers. Euphytica 133, 243-252.
24. Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A., Allard, R.W. (1984). Ribosomal 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, 8014-8019.
25. Saini, N., Jain, N., Jain, S., Jain, R.K. (2004). Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. Euphytica.140, 133-146.
26. Salem, K.F. and Sallam, A. (2016). Analysis of population structure and genetic diversity of Egyptian and exotic rice (*Oryza sativa* L.) genotypes. Comptes Rendus Biologie 1, 339(1):1-9.
27. Siwach, P., Jain, S., Saini, N., Chowdhury, V.K.,Jain, R.K. (2004). Allelic diversity among Basmati and Non-Basmati long-grain *Indica* rice varieties using microsatellite markers. Journal of Plant Biochemistry and Biotechnology 13, 25-32.
28. Temnykh, S., Park, W.D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y.G., Ishii, T., McCouch, S.R. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theoretical and Applied Genetics100, 697-712.