RESEARCH ARTICLE

GENETIC DIVERSITY ANALYSIS IN RICE (*ORYZA SATIVA* L.) LANDRACES OF NORTH EAST INDIA USING MORPHOLOGICAL AND RAPD MARKERS

Rinya Punyo and H.P. Chaturvedi*

Department of Genetics and Plant Breeding, Nagaland University, SASRD, Medziphema- 797106, Nagaland
Email: hpchaturvedi68@gmail.com

Received-06.07.2020, Revised-25.07.2020

Abstract: Morphological and molecular characterizations of sixteen rice landraces of North East India were studied to observe genetic diversity and identification of superior genotypes for crop improvement program. Based on the relative magnitude of D^2 values, 16 genotypes were grouped into 3 clusters. Cluster I had the highest number of genotypes (9 genotypes), cluster II had 4 genotypes and cluster III had 3 genotypes. Cluster I showed the highest mean for plant height at maturity, cluster II for 1000 grain weight and cluster III for number of effective tillers/plant, number of filled grains/panicle and harvest index. 1000-grain weight showed the highest contribution towards divergence followed by plant height. In RAPD analysis some of the primers showed 100% polymorphism viz., AA-10, AA-14, OPA-02, OPA-04, OPB-03 and PRIMER-33. Overall percentage polymorphism revealed by RAPD primer was 77.75%. Dendrogram generated from the UPGMA cluster analysis divides the lines into two main clusters- cluster A and cluster B. Cluster A consist of 15 germplasm and Cluster B consist of 1 germplasm. Cluster A is further divided into two sub-clusters, 'a' and 'b'. Sub-cluster 'a' consist of 2 germplasm and sub-cluster 'b' consist of 13 germplasm. Principal coordinate analysis (PCoA) obtained for RAPD is in complete support of the conclusion drawn from the cluster analysis.

Keywords: Genetic diversity, Morphology, Rice

INTRODUCTION

Rice (*Oryza sativa* L.) is a major staple food crop in the world. It is planted on approximately 11 percent of earth's cultivated land area (Anis *et al*., 2016). Almost 80% of the energy needs are derived from rice by over two billion people in Asia (Pragnya *et al*., 2018). Rice is one of the most important food crops, feeds more than 60 percent population of India. Rice is the principal food grain crop of the North Eastern states. It is the household food and nutritional security of the North-Eastern states. It occupies 3.51 million hectares which accounts for more than 80 percent of the total cultivated area of the region and 7.8 percent of the total rice in India while its share in national rice production is only 5.9 percent (Ngachan *et al*., 2011). The total rice production of the NE region is estimated to be around 5.50 million tonnes with average productivity of 1.57 t/ha, which is much below the national average of 2.08 t/ha (Pattanayak *et al*., 2006). The NE region is considered to be one of the hot pockets of the rice genetic resources in the world and a potential rice growing region with extremely diverse rice growing conditions as compared to other parts of the world (Ngachan *et al*., 2011). In general, diverse landraces traditionally are considered important for future food security due to their ability to sustain in changing climate (Huang *et al*., 2012; Pusadee *et al*., 2009 and Doebley *et al*., 2006). Therefore, the present study was carried out to analyze the genetic diversity in rice genotypes of North East India to select the potential parents for breeding program.

MATERIALS AND METHODS

The present experiment was carried out in the experimental farm of SASRD, NU, Medziphema campus using randomized block design with three replications. The experimental material comprises of sixteen genotypes of rice (Table- 1). The main objective of the experiment was to study the genetic diversity in available germplasm and to identify suitable genotypes. Rice seeds were sown on 22 June, 2018. Two- three seeds per hole were dibbled at 20 cm inter-row and 10 cm inter-plant distance. A population of 30 hills per plot was maintained. All the cultural practices were followed for raising a good crop. The data were recorded on five randomly sampled plants in each plot for 6 characters viz., Plant height at maturity, number of filled grains/panicle, number of effective tillers/plant, harvest index, 1000-grain weight and yield per plant. The mean values were subjected to statistical analysis to work out analysis of variance for all the characters as suggested Panse and Sukhatme (1957). The genetic divergence was estimated using Mahalanobis's D^2 statistics (Mahalanobis, 1936). All the genotypes were grouped into clusters on the basis of D^2 values by using Tocher's method (Rao, 1952).

*Corresponding Author

Table 1.

Sl. No.	Treatment	Local name	Place of collection
1	T_1	PYAPING LANCHA MIIPYA	ARUNACHAL PRADESH
2	T_2	JII PYAPING EMMO	ARUNACHAL PRADESH
3	T_3	KEMENYA KETHENGU KEMERIE-U NHALIEZHA	NAGALAND
4	T_4	TIIPE MIIPYA	ARUNACHAL PRADESH
5	T_5	KOGYA MIIPYA	ARUNACHAL PRADESH
6	T_6	PYATII MIIPYA	ARUNACHAL PRADESH
7	T_7	TSAMHO TSIA NHALIELHA	NAGALAND
8	T_8	INGLONGKIRI	ASSAM
9	T_9	HATCHA	ASSAM
10	T_{10}	PULU PYAPING EMMO	ARUNACHAL PRADESH
11	T_{11}	MISANG EMMO	ARUNACHAL PRADESH
12	T_{12}	AREY EMMO	ARUNACHAL PRADESH
13	T_{13}	AMPU EMMO	ARUNACHAL PRADESH
14	T_{14}	ELLANG EMMO	ARUNACHAL PRADESH
15	T_{15}	AMPU HATI EMMO	ARUNACHAL PRADSEH
16	T_{16}	TIMPI MIIPYA	ARUNACHAL PRADESH

For molecular diversity analysis the rice germplasms were planted at CSIR-NEIST, Jorhat, Assam. Young leaves were collected in bulk, washed, and then lyophilize for 48 hrs at -40^{o}c and stored at -20^{o}c till the isolation of DNA. DNA isolation was carried out by CTAB method. UV-absorption spectrophotometer was used to estimate the purity and concentration of the isolated genomic DNA samples. 260 nm and 280 nm ratio of absorbance was used as an indicator of DNA purity. A ratio of 1.8 is generally accepted as pure DNA.

A total of 40 RAPD primers were tested out of which 21 primers showed positive result (Table- 2). For the isolation, stock DNA is to be diluted to 15ng/µl. For RAPD the PCR reaction mixture was standardised by adding 3µl of diluted DNA,10µl of Hi- Chrome PCR Master Mix, 1.6µl of primer in the ratio 1:9 (IDT, India) to a total volume of 20 µl and the final volume was adjusted by adding double distilled sterilized water.

PCR setting:

For RAPD the standardized reaction programme was set at 94°C for 3 min followed by 35 cycles of 94°C for 54 s, primer annealing 45°C for 45 s, 2 min elongation at 72°C and a final extension of 72°C for 10 min as given below (Table- 3).

Table 2. List of RAPD markers and their sequence

Sl. No.	RAPD primers	Primer sequence
1	A-03	AGTCAGCCAC
2	AA-10	TGGTCGGGTG
3	AA-14	AACGGGCCAA
4	OPA-01	CAGGCCCTTC
5	OPA-02	TGCCGAGCTG
6	OPA-03	AGTCAGCCAC
7	OPA-04	AATCGGGCTG
8	OPA-07	GAAACGGGTG
9	OPA-13	CAGCACCCAC
10	OPB-03	CAGCACCCAC
11	OPB-04	GGACTGGAGT
12	OPB-06	TGCTCTGCCC
13	OPB-07	GGTGACGCAG
14	OPB-09	TGGGGGACTC
15	OPB-10	CTGCTGGGTAC
16	OPJ-06	TCGTTCCGCA
17	PRIMER-1	AATCGGGCTG
18	PRIMER-31	AAGCCTCGTC
19	PRIMER-33	GACGGATCAG
20	PRIMER-45	TGCGGCTGAG
21	PRIMER-59	CACAGACACC

Table 3. Configuration of PCR for RAPD

STEP	PROCESS	TEMPERATURE	TIME	CYCLES
STEP I	Initial denaturation	94^0c	3 minutes	35 cycles
	Denaturation	94^0c	54 seconds	
STEP II	Primer annealing	45^0c	45 seconds	
	Elongation	72^0c	2 minutes	
STEP III	Final elongation	72^0c	10 minutes	

The amplified product was visualized under UV light and documented by using a gel documentation system.

For the analysis of polymorphism in RAPD markers different parameters viz., total number of bands obtained for RAPD marker, number of polymorphic bands and percentage of polymorphic bands were calculated. The bands obtained from the markers were scored according to the presence as 1 and absence as 0. Polymorphism Information Content (PIC) was calculated to find the effectiveness of different RAPD primers. PIC value was calculated following the formula of Botstein *et al*. (1980) as—

$PIC=1 - \Sigma fi^2$ where f is the frequency of *i*th allele.

POPGENE software (version 1.32) by Yeh *et al*. (2000) was used to determine different genetic parameters such as- number of alleles per locus (Na), effective number of alleles (Ne), Nei's gene diversity (He), Shannon's diversity index (I). To determine genetic similarity dendrogram was constructed using NTSYS software programme version 2.1 (Rohlf, 2000). Principal coordinate analysis (PCoA) was done using Rohlf NTSYS Software to further differentiate the groups.

RESULT AND DISCUSSION

Analysis of variance revealed significant variation among the 16 genotypes for all the characters. The genetic divergence was estimated by utilizing Mahalanobis D^2 statistic as described by Rao (1952). Using this technique, all the genotypes were grouped into 3 different clusters based on the relative magnitude of D^2 values (Table-4). Cluster I had highest number of genotypes (9) followed by cluster II (4) and cluster III (3).

Table 4. Clustering pattern of 16 genotypes of rice on the basis of Genetic Divergence

Cluster number	Number of Genotypes	Genotypes
Cluster I	9	PYAPING LANCHA MIIPYA, TIIPE MIIPYA, KOGYA MIIPYA, KEMENYA KETHENGU, PULU PAYING EMMO, MISANG EMMO, AREY EMMO, AMPU EMMO, AMPU HATI EMMO
Cluster II	4	JII PAYING EMMO, PYATII MIIPYA, TSAMHO TSIA NHALIELHA, ELLANG EMMO
Cluster III	3	INGLONGKIRI, HATCHA, TIMPI MIIPYA

The estimates of intra and inter cluster distances have been presented in Table- 5. The maximum intra-cluster distance was observed in cluster III (33.81) and minimum in cluster I (23.59). The inter-cluster distance was found to be highest between clusters II and cluster III (111.05). The minimum inter-cluster was observed between clusters I and cluster II (57.83). The mean performances of all the characters are shown in Table- 6. Cluster I showed high mean value for plant height at maturity (59.38), cluster II showed high mean value for 1000-grain weight (26.60) and cluster III for number of effective tillers/panicle (2.74), number of filled grains/panicle (34.05), harvest index (6.82) and yield/plant (2.40). The relative contribution of each character to the total divergence is given in Table-7. 1000-grain weight contributed maximum towards total divergence (97.5%).

D^2 statistics is a useful tool to measure genetic divergence among genotypes in any crop. In the present study, the genotypes of cluster II and cluster III were found to be genetically diverse because highest inter-cluster distance was observed between these two clusters. Hence, genotypes from these two clusters i.e., cluster II and cluster III would be utilized in hybridization program to achieve greater variability in the segregating generations. To achieve higher variability and high heterotic effect,

Chaturvedi and Maurya (2005) also recommended that parents should be selected from two clusters having wider inter cluster distance. The minimum inter-cluster distance was observed between clusters I and cluster II, indicating close relationship between the genotypes belonging to these clusters.

The genotypes of cluster I recorded the superior performance for plant height at maturity while the genotypes of cluster II recorded the superior performance for 1000-grain weight and cluster III for number of filled grains/panicle, number of effective tillers per plant, harvest index and yield/plant. Cluster based mean estimations are useful in targeting the genotypes for breeding programme, as they prevent the tedious efforts of screening the inferior.

Table 5. Average Inter and Intra cluster distance of 16 Rice genotypes

Clusters	**Cluster I**	**Cluster II**	**Cluster III**
Cluster I	**556.69** (23.59)	**3344.17** (57.83)	**3639.09** (60.32)
Cluster II		**633.95** (25.18)	**12332.95** (111.05)
Cluster III			**1143.35** (33.81)

*D values are in Parenthesis

Table 6. Cluster wise mean values of 16 Rice Genotypes

Characters	**Plant height at maturity**	**Number of effective tillers/panicle**	**Number of filled grains/panicle**	**Harvest index**	**1000 grain weight**	**Yield/plant**
Cluster I	59.38	2.27	16.74	3.96	23.31	1.26
Cluster II	54.34	2.35	14.79	4.36	26.60	1.49
Cluster III	50.28	2.74	34.05	6.82	20.01	2.40

Table 7. Contribution of each Character towards Divergence

Sl. Number	**Characters**	**Times Ranked 1st**	**Contribution %**
1	**Plant height at maturity**	3	2.5%
2	**Number of effective tillers/ panicle**	-	-
3	**Number of filled grains/ panicle**	-	-
4	**Harvest index**	-	-
5	**1000 grain weight**	117	97.5%
6	**Yield/plant**	-	-

germplasm lines. These traits hold great promise as parents for obtaining promising elite lines through hybridization and to create further variability for these characters (Mishra and Pravin, 2004). The relative contribution of each character to the total divergence estimated by D^2 analysis indicated that 1000-grain weight contributed the maximum towards the genetic divergence. Chakma *et al*. (2012) also reported highest divergence for 1000-grain weight.

RAPD banding pattern and PIC

A total of 73 bands were produced by 21 RAPD primers, out of which 59 were polymorphic for the entire data set. The number of polymorphic bands ranged from 1 (RAPD A-03 and primer 31) to a maximum of 5 (RAPD OPA-04, OPB-06 and OPB-03). Primer AA-10, AA-14, OPA-02, OPA-04, OPA-07, OPB-03 and PRIMER-33 resulted in 100% polymorphism (Table- 8). Overall percentage polymorphism revealed by RAPD primers was 77.75%. PIC value of the primers ranged from 0.17 (Primer A-03) to 0.56 (OPB-03 and PRIMER-1).

Table 8. Primer sequence, polymorphic bands, percentage polymorphism and PIC value of RAPD

RAPD primers	Primer sequence	Number of polymorphic bands	Total number of bands	Percentage of polymorphic bands	PIC value
A-03	AGTCAGCCAC	1	2	50	0.17
AA-10	TGGTCGGGTG	2	2	100	0.49

AA-14	AACGGGCCAA	3	3	100	0.41
OPA-01	CAGGCCCTTC	3	5	60	0.23
OPA-02	TGCCGAGCTG	3	3	100	0.47
OPA-03	AGTCAGCCAC	2	3	66.6	0.32
OPA-04	AATCGGGCTG	5	5	100	0.53
OPA-07	GAAACGGGTG	3	3	100	0.54
OPA-13	CAGCACCCAC	2	3	60	0.29
OPB-03	CAGCACCCAC	5	5	100	0.56
OPB-04	GGACTGGAGT	3	4	75	0.42
OPB-06	TGCTCTGCCC	5	6	83.33	0.34
OPB-07	GGTGACGCAG	2	3	66.6	0.33
OPB-09	TGGGGGACTC	2	3	66.6	0.38
OPB-10	CTGCTGGGTAC	4	5	80	0.42
OPJ-06	TCGTTCCGCA	3	4	75	0.41
PRIMER-1	AATCGGGCTG	2	3	66.6	0.56
PRIMER-31	AAGCCTCGTC	1	2	50	0.26
PRIMER-33	GACGGATCAG	3	3	100	0.43
PRIMER-45	TGCGGCTGAG	2	3	66.6	0.35
PRIMER-59	CACAGACACC	2	3	66.6	0.52
TOTAL		59	73	-	-
MINIMUM		1	2	50	0.17
MAXIMUM		5	6	100	0.56
AVERAGE		2.8	3.4	77.75	0.40

Genetic variability in Rice germplasm

Genetic parameters *viz*- observed number of alleles, effective number of alleles, Nei's gene diversity, Shannon's diversity index are shown in Table- 9.

Table 9. Analysis of genetic parameters based on RAPD markers

Marker	Mean Na	Mean Ne	Mean He	Mean I
RAPD	1.79	1.47	0.27	0.41

Observed number of alleles (Na), effective number of alleles (Ne), Nei's gene diversity or expected heterozygosity (He), Shannon's diversity index (I)

Cluster analysis and principal coordinate analysis

Dendrogram generated from the UPGMA cluster analysis divides the lines into two main clusters- cluster A and B. Cluster A comprises of 15 germplasm and Cluster B consist of 1 germplasm. Cluster A is further divided into two sub-clusters, a and b. Sub-cluster 'a' consist of 2 germplasm and sub-cluster 'b' consist of 13 germplasm (Fig-1).

Principal coordinate analysis (PCoA) carried out for RAPD is shown in Fig-2. The PCoA plot obtained for RAPD marker is in complete support with the conclusion drawn from the cluster analysis.

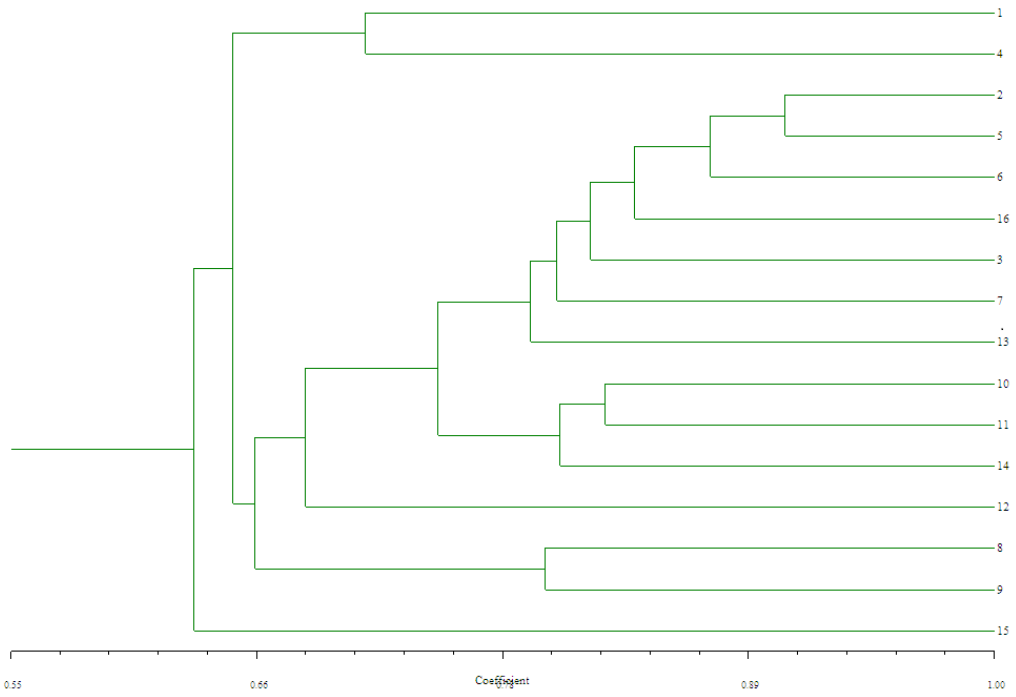

Fig. 1. Dendrogram generated from RAPD

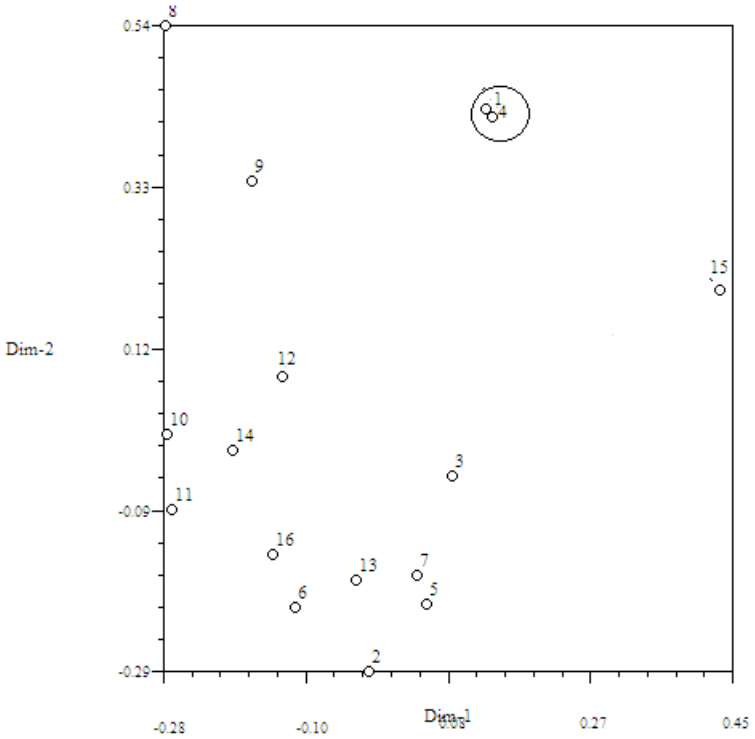

Fig. 2. PCoA analysis of Rice germplasm based on RAPD

Diversity in plant genetic resources (PGR) provides opportunity for plant breeders to develop new and improved cultivars with desirable characteristics. Germplasm identification is important for conservation and potential utilization of plant genetic resource. The molecular marker is a useful tool for assessing genetic variations and resolving cultivar identities. In the present study molecular characterization of 16 rice germplasms were done by using 21 RAPD primers. The study showed that the RAPD markers were effective enough to differentiate the population. The effectiveness of the markers was evaluated with the help of PIC. The PIC value ranged from 0.17 to 0.56. The percentage of polymorphism revealed by the RAPD markers (77.75%) is high compared to the previous studies, for example

Kanawapee *et al*., 2011 and Kiani *et al*. 2011 showed 68.94% and 67.35% polymorphism respectively among rice cultivars. The values of mean number of alleles (Na),mean effective number of alleles (Ne), mean Nei's gene diversity (He) and mean Shannon's index (I) were calculated as 1.79, 1.47, 0.27 and 0.41 respectively. On the basis of the observations in the present study, it is suggested that RAPD analysis can be effectively utilized for the purpose of varietal identification and detection of genetic variability through DNA polymorphism (Kanawapee *et al*. 2011 and Ren *et al.,*2003).

Cluster analysis drawn using UPGMA method showed distinct separation within the germplasm. It is also in complete agreement with the data obtained from the PCoA plot.

REFERENCES

Anis, G., EL Sabagh, A., Ghareb, A. and EL-Rewainy, I. (2016). Evaluation of promising lines in rice (*Oryza sativa* L.) to agronomic and genetic performance under Egyptian conditions. *International Journal of Agronomy and Agricultural Research*. **8**(3): 52-57.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet*. 32: 314-331.

Chakma, S.P., Huq, H., Mahmud, F. and Husna, A. (2012). Genetic Diversity Analysis in Rice (*Oryza sativa* L.). *Bangladesh Journal of Plant Breeding and Genetics* **25**(1). 31-39.

Chaturvedi, H. P. and Maurya, D. M. (2005). Genetic divergence analysis in rice (*Oryza sativa* L.). Advances Pl. Sciences. 18(1):349-353.

Doebley, J.F., Gaut, B.S. and Smith, B.D. (2006). The molecular genetics of crop domestication. Cell **127**:1309-1321.

Huang, X., Kurata, N., Wei, X. *et al.* (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**:497–501.

Kanawapee, N., Sanitchon, J., Srihaban, P. and Theerakulpisut, P. (2011). Genetic diversity analysis of rice cultivars (Oryza sativa L.) differing in salinity tolerance based on RAPD and SSR markers. *Electronic Journal of Biotechnology*. **14**(6): DOI: 10.2225/vol 14-Issue6.

Kiani, G. (2011). Genetic diversity analysis of Iranian improved rice cultivars through RAPD markers. *Not. Sci. Biol*. **3**(3): 135-139.

Mahalanobis, P.C. (1936). On the generalized distance in statistics. *Proc. Natl. Inst. Sci*. India. (2): 49-55.

Mishra, L.K. and Pravin, J. (2004). Variability and genetic diversity in rice (Oryza sativaL.). Mysore J. Agric.Sci. 38: 367-375.

Ngachan, S.V., Mohanty, A.K. and Pattanayak, A. (2011). Status Paper on North East India. ICAR, Research Complex for NEH Region.

Panse, V. G. and Sukhatme, P. V. (1957). Statistical methods for agricultural workers. Indian Council of Agricultural Research, New Delhi.

Pattanayak, A., Bujarbaruah, K.M., Sharma, V.P., Ngachan, S.V., Dhiman, K.R., Munda, G.C., Azad Thakur, N.S. Satpathy, K.K., Dutta, K.K., Pakash, N., Devi, P. and Vishwakarma, A.K. (2006). Step towards optimizing rice production in North East India. *J. Research Bulletin*. No-21, ICAR Research Complex for North East Hill Region, Barapani.

Prangya, K., Radha, K. K.V., Subba, R. L.V., and Suneetha, K. (2018). Studies on genetic divergence analysis in soft rice *(Oryza sativa L.)* genotypes. *International Journal for Pure and Applied Bioscience* **6**(2): 968-975.

Pusadee, T., Jamjod, S., Chiang, Y.C., Rerkasem, B. and Schaal, B.A. (2009). Genetic structure and isolation by distance in a landrace of Thai rice. Proc Natl Acad Sci USA **106**:13880–13885.

Rao, C.R. (1952). Advance statistical methods in Biometrical Research. John Wiley and Sons, Increased. New York.

Ren, F.; LU, B-R.; LI, S.; Huang, J. and Zhu, Y. (2003). A comparative study of genetic relationships among the AA-genome *Oryza* species using RAPD and SSR markers. *TAG Theoretical and Applied Genetics*, vol. 108, no. 1, p. 113-120.

Rohlf, F.J. (2000). NTSYS-pc: numerical taxonomy and multivariate analysis system version 2.1. Exeter Software. Setauket, NY.

Yeh, F.C., Yang, R.C., 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.