Characterization and Heterotic Grouping of Traditional Assam rice (*Oryza sativa* L.)

Praveen Kumar¹, Debojit Sarma², and ³Mainu Hazarika

¹PhD Scholar, Department of Plant Breeding and Genetics, Assam Agricultural University, Jorhat-785013, Assam, India

²Professor, Department of Plant Breeding and Genetics, Assam Agricultural University, Jorhat-785013, Assam, India

³PhD Scholar, Department of Horticulture, Assam Agricultural University, Jorhat-785013, Assam, India

Abstract: Parents of heterotic hybrids are derived from different heterotic groups with high genetic divergence. Classification of traditional Assam rice germplasm in divergent pools will be advantageous to maximize the heterosis and thereby to ensure food security. In the present investigation, a group of 60 upland rice genotypes were characterized using 53 polymorphic simple sequence repeats (SSR) markers out of 83 molecular markers. The genetic divergence study using unweighted Neighbour-joining (UNJ) method clustered the 60 genotypes into 3 major clusters. The eleven most divergent genotypes identified were crossed in half diallel fashion to determine the mid-parent and better-parent heterosis values for the objective of heterotic grouping. No correlation between heterosis and genetic distance can be attributable to the use of a subset of markers not linked to yield or concerned. In genetic distance based heterotic grouping, the intra-group hybrids were recorded a higher frequency of crosses, grain yield per plant, specific combining ability effect, mid parent heterosis, better parent heterosis and standard parent heterosis value than those of inter-group hybrids. Overall, an extensive choice of parents with attractive traits constellation leading to increased yield of the hybrids for much better complementation must be stressed along with a substantial hereditary distance for augmentation of yield heterosis.

Keywords: Genetic distance, heterotic grouping, heterosis, SSR marker

Introduction

The global population comprises over 7 billion people, and the global population clock is recording continuous growth (Borem et al. 2014). Rice is the second most important crop in the world, feeding more than one half of the world population (Sharma et al. 2020). Above 90% of rice is consumed by Asians. Today the population growth is the alarming rate, so hybrid rice is a viable and practical approach for feeding tremendously increasing population of India as well as the world. Rice provides almost 15% of worldwide per capita protein requirements (IRRI 2002). The achievement of the hybrid rice that resulted from the exploitation of heterosis is based on the genetic divergence of germplasm, mostly on geographic change for the three-line hybrid rice and sub-specific genetic divergence for the two-line hybrid rice. North-East (NE) India is a traditional home for a vast array of rice germplasm, and high genetic diversity is expected for the landraces belonging to this region (Sahoo et al. 2019). Because of its unique eco-geographical features and broad ethnic diversity, it possesses at least 10,000 indigenous rice landraces that cultivated under lowland, upland, and deep water conditions (Hore et al. 2005).

Rice occupies 4.58 million hectares which are 75% of the total cultivated area of the region (GOI 2015). The rice landraces are grown in diverse ecosystems spreading from the high altitude of Arunachal Pradesh, the food-prone regions of Assam, and rainfed, irrigated, upland, steep terraces and deepwater, Jhum and *tilla* land ecologies. Of the various classes of rice cultivated in Assam, upland rice cultivars of North East India are directly sown in the fields during March-April and harvested in June-July known agronomically as "*aus/ahu*" rice (Travis et al. 2015). Genetic diversity among landraces and their genetic relationship is helping in conservation and parental selection in the future breeding programme. Identification of landraces with a high level of genetic variation will be a valuable resource for broadening the genetic base. Since it allows superior alleles to be identified for improved quantitative traits. On the other hand, DNA-based molecular markers are persistent, repeatable, steady, vigorous and profoundly dependable (Virk et al. 2000). Among the accessible DNA markers, simple sequence repeat (SSR) is suitable because of their multi-allelic nature, high reproducibility, co-dominant nature, plenitude and genome-wide coverage (Temnykh et al. 2000), which differ in the level of polymorphism relying upon their area in the coding or non-coding portions, nature of their recurrent themes and the genome-wide coverage. Along these lines, a perfect arrangement of SSR markers giving genome-wide inclusion encourages the measure of the hereditary variation, which provides an intense, unambiguous molecular depiction of rice cultivars/landraces. Previously, numerous scientists have reported the evaluation of the descend varieties of Indian rice germplasm (Sivaranjani et al. 2010; Vanniarajan et al. 2012).

Parents of heterotic hybrids are usually derived from different heterotic groups with high genetic divergence. Grouping of germplasm in divergent pools is advantageous to maximize the expected heterosis (Reif et al. 2005). Heterotic groups for hybrid crops could be determined by marker-based groups as studied in maize (Choukan et al. 2006; Lu et al. 2009; Romay et al. 2013; Xie et al. 2008), chickpea (Sant et al. 1999), rice (Kwon et al. 2002), alfalfa (Riday et al. 2003; Tams et al. 2006) and in maize (Geleta et al. 2004) confirmed the non-significant correlation of genetic distance with heterosis. So, the genetic distance may not be as such reliable for rice hybrid breeding programme and predicting the heterosis at the DNA level. Due to inadequate information on the association of functional molecular markers and yield, the evidence derived from molecular markers at this time is limited to the use of assigning parents into germplasm group or heterotic groups, and to provide a general guideline of avoiding

heterotic groups from blending during parent breeding. Full detailed information about functional markers related to yield heterosis may provide precise direction to upsurge the breeding efficiency of rice.

Material and method

Experimental materials

A collection of 60 rice genotypes consisting of landraces, maintainers of wild abortive cytoplasmic male sterile (WA-CMS) lines and high yielding varieties from Assam, India were used in the present study. Pedigree and origin of the genotypes are presented in Table 1. The latitude, longitude and altitude of Jorhat are 26°44'N, 94°10'E and 91 m. above mean sea level respectively. All the molecular work, including DNA extraction, PCR and gel electrophoresis, were performed in the marker laboratory of the Department of Agricultural Biotechnology, Assam Agricultural University, Jorhat. The total genomic DNA from each of the 60 rice genotypes was extracted from 5 g leaves of 21 days old seedlings following the protocol of (Dellaporta et al. 1983) with minor modifications.

Molecular data analysis

Genetic distance (GD) between each pair of parents was measured (Cavalli-Sforza and Edwards 1967), using PowerMarker version 3.25 (Liu and Muse 2005). Based on Cavalli-Sforza and Edwards chord distances, a dendrogram was constructed illustrating the genetic relationship among the rice genotypes using unweighted Neighbour-joining (UNJ) method as proposed by (Gascuel 1997), which uses a criterion of weighted average, in DARwin 6 (Perrier and Jacquemoud-Collet 2006). The number of alleles per locus, major allele frequency, gene diversity, heterozygosity and polymorphism information content (PIC) values was calculated using PowerMarker version 3.25 (Liu and Muse 2005).

Cavalli-Sforza chord distance

Cavalli-Sforza and Edwards 1967 proposed the chord distance assuming that the genetic differences arise due to genetic drift only. The chord distance between two populations represented on the surface of a multidimensional hypersphere using allele frequencies at the j^{th} locus is given by

$$D_{CH} = \frac{2}{\pi m} \sum_{j=1}^m \sqrt{2 \left(1 - \sum_{i=1}^{a_j} p_{ij} q_{ij} \right)}$$

Let p_{ij} and q_{ij} be the frequencies of an i^{th} allele at the j^{th} locus in populations X and Y , respectively and a_j is the number of alleles at the j^{th} locus, and m is the number of loci examined? The binary data were used to generate a dissimilarity matrix and genetic diversity using chord distance in PowerMarker version 3.25 (Liu and Muse 2005).

Gene diversity (He)

The gene diversity of expected heterozygosity (He) is defined as the probability that two randomly chosen alleles from the population are different. Gene diversity (Liu and Muse 2005) is defined as follows:

$$He = 1 - \sum_{i=1}^m p_i^2$$

Where, m is the number of alleles in a gene locus and p is the i^{th} allele frequency.

Polymorphic information content (PIC)

The Polymorphic information content (PIC) values provide an estimate of the discriminatory power of the locus or loci by taking into account not only the number of alleles that are expressed but also the relative frequency of those alleles. The informative potential of a marker is high if the PIC value is more than 0.5, moderate if the PIC is between 0.5 and 0.25, and only slightly informative if the PIC value is below 0.25 (Ditta et al. 2018). The polymorphic information content (PIC) for each locus (Botstein et al. 1980) is defined as follows:

$$PIC_i = 1 - \sum_{j=1}^m p_i^2 - \sum_{i=1}^{m-1} \sum_{j=i+1}^m 2p_i^2 p_j^2$$

Where, p_i and p_j represent the frequencies of the i^{th} and the j^{th} alleles, respectively and m is the number of alleles in a locus.

Morphological Characters

All the observations except seedling establishment, days to panicle initiation, 50% flowering and maturity were based on 3 random competitive plants. The panicles of the main culms of the sampled plants were cut just below the neck and were separately bagged. Shoot dry weights inclusive of grains for the sample plants were recorded after drying in a hot air oven at 70°C until constant weight. Initial weights of the harvested shoots with intact grains and the separated grains were noted. The filled grains were separated from the remaining shoot portion of each sampled plant before drying to enable moisture determination of the shoots inclusive of grains and the grains alone. The characters namely average panicle weight, straw yield, grain yield and biological yield were reported at 12 per cent moisture content. The observations on flag leaf length/ width were taken at booting stage and on culm length, productive tillers, panicle length was recorded during physiological maturity stage in the standing crop.

Seedling height at 21 DAS (SH) the average length of five random seedlings (shoot portion) was measured in cm at 21 days after sowing (DAS).

Leaf number at 21 DAS (LN) the total number of leaves of five random seedlings were counted at 21 DAS and the average was taken to get leaf number per seedling.

The seedling establishment at 7 DAT (SEST) the seedling establishment was calculated as the ratio of the number of seedlings survived at 7 days after transplanting (DAT) to the total number of seedlings transplanted and expressed in percentage.

Days to panicle initiation [DPI] days to panicle initiation was recorded from the date of sowing to the date of first panicle emergence in each single row plot as a single observation.

Days to 50 per cent flowering [DFF] the number of days taken from sowing to 50% of the plants in each row with emerged panicles was recorded in days on a plot basis.

Flag leaf area (cm²) [FLA] the flag leaf area was calculated at booting stage using the formula given by Yoshida (1981) as follows: Leaf area (cm²) = 0.75 × length (cm) × width (cm)

Days to maturity [DM] it was recorded as the number of days taken from sowing to the time when more than 80 per cent of the grains on the panicles of 80 per cent of the plants in a row turned fully yellow.

Culm height (cm) [CH] the height of the main culms of the sampled plants were measured in cm from the ground level to the base of the panicle and the average worked out.

The number of productive tillers per plant [PT] the average number of panicle bearing tillers of the sampled plants per row was recorded at harvest stage.

Average panicle weight (g) (APW) the average weight of the panicles borne on main culms of the sampled plants were recorded in gram before counting the total number of grains.

Panicle length [PL] the average length of the panicles borne on the main culms of five random plants was measured in cm from the panicle neck to the tip of the panicle.

The number of filled grains per panicle [GP] the number of filled grains in the panicles borne on the main culms of the sampled plants was counted and the average worked out.

Spikelet fertility (%) [SF] the spikelet fertility was calculated as the ratio of the number of filled grains to the total number of grains per panicle expressed in percentage.

Straw yield per plant (g) [SYP] the average weight of the shoot portions excluding the grains for the sampled plants was reported as straw yield per plant in gram.

Grain yield per plant (g) [GYP] the average weight of the filled grains of the sampled plants was expressed as grain yield per plant in gram.

Biological yield per plant (g) [BYP] it was calculated as the sum of straw yield and grain yield of each sampled plant and the average expressed in gram.

Harvest index (%) [HI] the harvest index was calculated in percentage as the proportion of grain yield to the biological yield:

$$\text{Harvest index (\%)} = \frac{\text{Grain yield per plant}}{\text{Biological yield per plant}} \times 100$$

Gel consistency [GC] test was conducted by the method described by (Cogampang et al. 1973). The method separates high amylose rice into hard gel consistency (26-40 mm), medium gel consistency (41-60 mm) and soft gel consistency (61-100 mm) types.

Amylose content (%) [AC] amylose content was determined by using the method described by Dela Cruz and Khush (2000). Rice varieties were grouped based on their amylose content into waxy(0-2%), very low (3-9%), low (10-19%), intermediate (20-25%) and high (>25%)(Kumar and Khush 1986).

Heterosis Analysis: The estimates of heterosis were calculated as a percentage increase or decrease of F_1 s over the mid parent ($\overline{MP}$), better parent ($\overline{BP}$) and standard parent ($\overline{SP}$) values following the method of (Turner 1953; Hayes et al. 1955).

Heterosis over mid parent (H_1) $H_{\overline{MP}}(\%) = \frac{\overline{F_1} - \overline{MP}}{\overline{MP}} \times 100$

Heterosis over better parent (H_2) $H_{\overline{BP}}(\%) = \frac{\overline{F_1} - \overline{BP}}{\overline{BP}} \times 100$

Heterosis over standard parent (H_3) $H_{\overline{SP}}(\%) = \frac{\overline{F_1} - \overline{SP}}{\overline{SP}} \times 100$

Where, F_1 = mean of F_1 , $\overline{MP}$ = mean of the two parents, $\overline{BP}$ = mean of the better parent and $\overline{SP}$ = mean of the standard parent.

the heterosis was tested by the least significant difference at 5% and 1% level of significance for error degree of freedom (m) as follows. For testing heterosis over mid parent $SE_d(\overline{MP}) =$

$$\sqrt{\frac{3M'_e}{2}}$$

For testing heterosis over better parent and standard parent $SE_d(\overline{BP}/\overline{SP}) = \sqrt{2M'_e}$

Critical difference $CD = SE_d \times t$ (at 5 and 1% level for error d.f.), heterosis was considered significant when $(F_1 - \overline{MP})$ or $(F_1 - \overline{BP})$ was higher than the critical difference.

Result and discussion

Molecular marker analysis

Polymorphism of SSR markers

Among the classes of repetitive DNA sequences that have shown acquiescent for PCR amplification, SSRs remains the perfect choice of markers (Jacob et al. 1991). SSR markers are valuable genetic markers because they detect a high level of allelic diversity, and are co-dominant, accessible and economically assayed by PCR (Weber and May 1989). They are easily automated (Westman and Kresovich 1997), abundance and have even genomic distribution and high level of polymorphism (Verma et al. 2019). It has a common polymorphism of at least 1.5 times higher than AFLP and RAPD markers (Mackill et al. 1996). SSRs are highly polymorphic, even between closely related lines (Gupta et al. 1999). The polymorphism in SSR could be due to changes in the SSR region itself caused by the expansion or contraction of SSR or interruption (Li et al. 2000). In the present investigation, a total of 83 SSR markers distributed throughout the 12 chromosomes were used to assess the extent of molecular diversity across 60 rice genotypes of Assam maintained in the Department of Plant Breeding and Genetics of Assam Agricultural University, Jorhat. Out of 83 SSR markers tested, 53 were determined polymorphic.

Summary statistics about the SSR markers

DNA bands were scored for DNA fingerprinting analysis with the molecular data generated using 53 SSR markers. Some gel pictures of the polymorphic markers (RM 293, RM 447, RM 429, RM 337, RM 245, RM 152 and RM 216) showed (Supplementary Figure 1). The major allele frequency, number of alleles per locus, gene diversity, heterozygosity and PIC values (Table 2) were calculated for each SSR marker using Power Marker v 3.25 (Liu and Muse 2005). All the genotypes were scored for the allelic weight of the SSR bands throughout all the 60 genotypes. DARwin 6 used an unweighted neighbour method to construct the dendrogram showing the distance-based interrelationship among the genotypes. For the phylogenetic tree, genetic distance was calculated as the chord distance in PowerMarker v 3.25 (Liu and Muse 2005).

Major allele frequency

The major allelic frequency revealed by the SSR markers across the 60 rice genotypes ranged from 0.400 to 0.983, with a mean of 0.701. Generally, the allele frequency of a maximum number of markers was below 0.95, indicating that they were all polymorphic. Choudhury et al. 2014, reported major allele frequency ranging from 0.50 for Assam to 0.99 for Tripura. Gour et al. 2017, reported the average of major allele frequency from 0.42 to 0.10 and Mvuyekure et al. 2018, reported 0.76; all these findings were quite similar to our results.

Number of alleles

The SSR markers have significantly superior allelic diversity of microsatellites (McCouch et al. 1997) and high numbers of alleles for rice microsatellite markers. In the present study, a total of 133 alleles were detected across 60 rice genotypes by 53 polymorphic SSR markers with an average of 2.5 alleles per locus. The number of alleles generated per locus by each marker ranged from 2 (RM296, RM1896, RM204, RM249, RM12, RM519, RM11, RM20, RM44, RM318, RM530, RM3614, RM5361, RM21, RM202, RM287, RM50, RM520, RM29,

RM60, RM261, RM-325, RM28519, RM127, RM171, RM178, RM216, RM222, RM252, RM3331, RM1352, RM15669 and RM438 to 5 (RM260). These numbers were quite comparable to 2.0-5.5 alleles per SSR locus (Cho et al. 2000) using a different set of rice germplasm. A range of 2-4 alleles per locus was reported (Gour et al. 2017; Ming et al. 2015; Mvuyekure et al. 2018; Verma et al. 2019). The average number of alleles was 6.72 (Ravi et al. 2003), 4.90 (Zeng 2004), 4.50 (Pervaiz et al. 2010), 8.00 (Zhao et al. 2009), 4.30 (Islam et al. 2011), 7.40 (Shamim et al. 2015) and 6.40 (Watanabe et al. 2016), which might be due to the use of a more diverse set of rice accessions in their study or due to the use of highly polymorphic markers.

Gene diversity

Gene diversity ranged from 0.033 in RM204 to 0.706 in RM334, with a mean of 0.387. Gour et al. 2017 and Mvuyekure et al. 2018, reported mean gene diversity value of 0.390 and 0.325, respectively, which were similar to present mean value. Watanabe et al. 2016, reported a gene diversity mean of 0.84, and Choudhury et al. 2014, obtained gene diversity ranging from 0.006 in Arunachal Pradesh to 0.50 in Manipur.

Observed heterozygosity

The observed heterozygosity ranged from 0.017 for RM287 to 0.983 for RM334, with a mean of 0.152. The majority of the SSR markers exhibited observed heterozygosity either zero or low value. Watanabe et al. 2016 and Gour et al. 2017, reported heterozygosity of 0.690 and 0.386, respectively. Choudhury et al. 2014, reported heterozygosity ranging from 0.002 in Nagaland to 0.420 in Mizoram. The present results suggested that the majority of rice germplasm were pure and completely homozygous for SSR markers which might be the result of the self-pollinated mode of reproduction of rice. The observed heterozygosity (0.152) was far lower than total expected heterozygosity (0.387) which further supported a low gene flow value for the majority of the loci, except RM 440, RM 19, RM 260, RM 206, RM 178, RM 293, RM 447 and RM 334.

Polymorphism information content (PIC)

In the present investigation, the mean PIC value for all the markers RM 293, RM 334 and RM 545 showed higher discriminatory power to distinguish genotypes due to its high PIC values of 0.645, 0.561 and 0.549, respectively. The primer RM 60 and RM 573 showed lower PIC values 0.090 and 0.064, respectively, suggesting less discriminatory power of these primers. A microsatellite marker having a PIC value of greater than 0.5 is highly informative (Botstein et al. 1980). The average PIC values of 0.33 (Singh et al. 2004), 0.32 (Ming et al. 2015), 0.32 (Gour et al. 2017) and 0.26 (Mvuyekure et al. 2018), reported previously. Pachauri et al. 2013, obtained the mean PIC value of 0.37 in sets of 14 improved varieties and 27 landraces of rice collected from different zones of seven Indian states, which were comparable to the present result. All these findings indicated that the rice accessions used in the present study had a broad genetic diversity. The average PIC values reported were 0.57 (Ravi et al. 2003), 0.57 (Zeng 2004), 0.66 (Zhao et al. 2009), 0.46 (Seetharam et al. 2009), 0.53 (Pervaiz et al. 2010), 0.603 (Islam et al. 2011), 0.57 (Shamim et al. 2015), 0.665 (Yadav et al. 2015), and 0.47 (Anh et al. 2018), respectively. These values were comparatively higher than the earlier reports, which might be due to the high diversity of rice accessions used in their studies or the use of highly polymorphic markers.

The dendrogram showing hierarchical horizontal clustering of 60 *Ahu* rice genotypes of Assam using unweighted neighbour-joining (UNJ) method based on chord distances estimated from 53 polymorphic SSR marker data is presented in Fig. 1. The dendrogram indicated dissimilarity values ranging from 0.111 to 0.695 with 3 major clusters. Cluster I had two sub-clusters I-A and I-B containing 18 and 3 genotypes, respectively. Similarly, cluster II was the largest cluster with 34 genotypes which consisted of two sub-clusters II-A having 16 genotypes and II-B with 18 genotypes. Cluster III contained 5 genotypes belonging to landraces only. All the genotypes of cluster I belonged to landraces except Nagina 22 which was a selection from Rajbhog. Cluster II-A was an admixture composed of both improved varieties and landraces. All the maintainer lines of WA CMS belonged to cluster II A. Bau Murali, and Krishna were the most distantly apart genotypes among all the 60 genotypes. The cluster composition is shown in Table 3.

Fig. 2. showed 53 SSR markers used for hierarchical clustering using the Unweighted Neighbour-joining (UNJ) method of 11 *Ahu* rice genotypes used in the present investigation. The grouping resulted by using the chord genetic distances among the 11 parental genotypes. All the genotypes included in each cluster represent a heterotic group. As shown in Table 4, three broad heterotic groups were identified using 53 polymorphic SSR markers. Group, I contained BorMekohiDhan and Mayamoti, group II comprised of Luit, Lachit, IR 58025B, IR 68888B, IR 68897B, IR 79156B and IR 80555B, group III included Suryamukhi and Lal Aus. All the maintainers of WA CMS belonged to group II along with Luit and Lachit. Landraces belonged to two different groups I and III. Thus, a fair degree of similarity was evident among genotypes of the same group. Also, the genotypes having a similar combining ability performance and the heterotic response would categorize them into one heterotic group (Melchinger and Gumber 1998).

Clustering based on genetic distances

The heterotic pattern increases the efficiency of hybrid development, inbred recycling, and population improvement. Recognition and determination of heterotic groups and patterns are essential for breeding hybrid varieties, as shown in various studies on rice (He et al. 2012; Xie et al. 2014). Inbred varieties are then often developed from crosses within heterotic groups. Promising hybrids result from crossing the inbred lines developed between different heterotic groups. Grouping of germplasm in divergent pools is advantageous to maximize the expected heterosis (Reif et al. 2005). The success of the hybrid rice that resulted from the utilization of heterosis depends on the genetic divergence of germplasm, basically on geographic divergence for three-line hybrid rice and sub-specific genetic divergence for two-line hybrid rice. Heterotic groups for hybrid crops could be determined by marker-based groups as studied in maize (Choukan et al. 2006; Romay et al. 2013; Xie et al. 2008), as well as in other crops (Tams et al. 2006). He et al. 2012, investigated the genetic diversity of hybrid rice parents developed at IRRI and evaluated with simple sequence repeats and single-nucleotide polymorphism markers where they confirmed that heterotic hybrids could be formed based on marker-based parent groups to increase the efficiency of hybrid rice breeding (Xie et al. 2014). Fig. 3. showed the frequency distribution of genetic distances (GDs) of the 60 rice genotypes and the selected 11 parents for the diallel crosses based on 53 SSR markers. The mean GD for the 60 genotypes was 0.507 and for the 11 parents was 0.475. This 11 parents' sample accounted for 83.3% of the allelic variation of the 60 genotypes and reasonably represents the population panel because of the high allelic coverage and similar cluster structure. This result was closely related to Wang et al. 2015.

Pooled analysis of variance (ANOVA)

The pooled analyses of variance (ANOVA) for the various traits evaluated under 40 and 60 kg N ha⁻¹ are presented in **Table 5.1 to 5.3**. The results revealed highly significant ($p < 0.01$) differences among the genotypes for all the traits, indicating that the material under investigation was diverse for the traits in question. The environment source of variation due to Nitrogen (N) doses showed significant differences ($p < 0.01$) for all the traits except only seedling establishment, culm length, panicle length, harvest index, amylose content. The genotype x environment (GE) interaction exhibited highly significant differences ($p < 0.01$) for all traits except seedling establishment, suggesting differential behaviour of the genotypes in the two N doses.

Mean comparison

Mean values of the environments (N-doses) averaged over the genotypes for the different traits showing significant environments source of variation are given in Table 5.4 and 5.5. The traits namely, days to panicle initiation (77.7), days to 50% flowering (83.6), flag leaf area (27.1 cm²) days to maturity (111.1), productive tillers plant⁻¹ (16.3), average panicle weight (3.3 g), grains panicle⁻¹ (116.1), spikelet fertility (78.0%), straw yield plant⁻¹ (62.7 g), grain yield plant⁻¹ (28.5 g), biological yield plant⁻¹ (91.0 g), gel consistency (49.8 mm) obtained for 60 kg N ha⁻¹ were significantly higher ($P < 0.05$) than the respective values of the traits at 40 kg N ha⁻¹. Thus, increasing the N-dose from 40 to 60 kg had a positive correlation with grain yield and its components traits. Increasing grain yield and biomass could be due to N-supply increase chlorophyll content, leaf area index, and nutrient uptake and utilization. Significant effect of increasing nitrogen rate on rice grain yield was also reported by (Bhuiya et al. 1989; Hussain et al. 1989; Islam et al. 1990; Ebaid and Ghanem 2000; El-Batal et al. 2004; Salem 2006; Wang et al. 2017; Adhikari et al. 2018). Adhikari et al. 2018; Liu et al. 2019, reported that nitrogen application increases the straw yields of transplanted rice. Wang et al. 2017, obtained that N fertilizer doses significantly increased the number of tillers in rice; late-emerging tillers usually produce lower yields compared with early emerging tillers. Mean performances of the parents and the hybrid groups classified as maintainers (M), improved varieties (IV) and landraces (LR) for the traits showing significant genotypic variation are given in **Table 5.6 and 5.7**. Among the two-hybrid groups involving WA CMS maintainers and the remaining parents, M×LR was the most productive for grain yield and other related traits, while LR×IV yielded the maximum with a concomitant increase in component traits. The hybrid group IV×IV was the least productive in terms of yield and most yield component traits, followed by LR×LR and M×M, suggesting that improved varieties were already selected for most desirable alleles for high yield and thus, hybridization among them resulted in less yield improvement, due to few allelic differences between them than crosses between landraces and improved varieties. Given the decrease in genetic diversity among the improved rice varieties due to shifting of the breeding goal from high yield to biotic and abiotic stresses, Singh et al. 2016, suggested broadening the genetic base by incorporating more diverse donor parents in the breeding programme for yield improvement in rice.

Range of mid-parent (H_{MP}) and better-parent heterosis (H_{BP})

The range of mid-parent heterosis for grains yield per plant was from -6.82** (LAL×BOR) to 31.11** (88B×LUI) (Table 6). The crosses, BOR×25B, LAL×55B, SUR×LAC and MAY×LAC showed significant positive mid-parent heterosis for grains yield per plant. Significant positive heterosis for grain yield per plant has been reported (Ghara et al. 2012; Kumar and Adilakshmi 2016; Latha et al. 2013; Thorat et al. 2017). Better parent heterosis for grain yield per plant ranged from -10.38** (LAL×BOR) to 29.20** (88B×LUI). The crosses, namely 88B×LUI, SUR×LAC, 88B×55B and MAY×LAC, exhibited significant positive better-parent heterosis for grains yield per plant. A similar finding was also reported by (Sahu et al. 2017).

Heterotic grouping based on genetic distances

Genetic distance based heterotic clustering of the 11 parental genotypes along with mean yield, combining ability and heterosis estimates (Table 7) revealed that the intra-group hybrid category G2×G2 contained the highest frequency of crosses (0.38) followed by inter-group hybrid category G1×G2 (0.25) and G2×G3 (0.25). G1×G2 (0.55) registered the highest genetic distance, followed by G1×G3 (0.48) and G2×G2 (0.40). G2×G3 (34.61) recorded the highest grain yield, followed by G3×G3 (33.76) G1×G1 (28.48). G3×G3 (7.03) showed the highest SUM-GCA (g plant-1) followed by G2×G3 (3.05) and G1×G3 (1.63). G1×G1 (5.02) showed the highest SCA (g plant-1) followed by G2×G3 (4.35) and G1×G2 (1.72).

The highest H_{MP} (%) was observed for G2×G3 (63.70) followed by G1×G1 (51.74) and G3×G3 (46.05). G2×G3 (46.15) exhibited the highest H_{BP} (%), followed by G1×G1 (43.56) and G3×G3 (36.00). G2×G3 (75.60) showed the highest H_{SP} (%), followed by G3×G3 (71.26) and G1×G1 (44.51). G2×G3 had the highest heterosis, grain yield and genetic distance; therefore, hybrids from this group are useful to produce superior hybrids, predicting the heterosis and formed a heterotic pattern. But the highest SCA was shown by G1×G1, an intra-group hybrid category. The lowest yielding hybrids and yield heterosis were evident in the crossing pattern of G1×G3.

The parental group G2 recorded the highest frequency of crosses (0.88), genetic distance (0.50), SCA (2.10), HMP (46.51%) and HBP (32.68%). G3 registered the highest grain yield (31.13), SUM-GCA (3.90) and HSP (57.93%). Mean of the inter-group hybrid categories for the frequency of crosses (0.42), genetic distance (0.53), grain yield (28.74), SUM-GCA (0.78), SCA (0.74), HMP (40.77%), HBP (28.20%) and HSP (45.81%) were 0.42, 0.53, 28.74, 0.78, 0.74, 40.77%, 28.20% and 45.81%, respectively. Mean of the intra-group hybrid categories were 0.58, 0.29, 29.59, 0.78, 1.59, 45.07%, 34.36% and 50.12% for the frequency of crosses, genetic distance, and grain yield, SUM-GCA, SCA, H_{MP} , H_{BP} and H_{SP} , respectively.

The inter-group parents were more diverse ($GD=0.53$) than the intra-group parents ($GD=0.29$), which was in tune with the findings of (Xie et al. 2014; Yingheng et al. 2018). However, the hybrids originating from inter-group parents do not always give high yield or yield heterosis (Wang et al. 2015). But some specific intra-group hybrids, G3×G3 in the present case, having low genetic distance produce a higher yield than the average yield of inter-group hybrids. The present study suggested that molecular markers might be useful for the grouping of parents and heterotic grouping. Based on this, an intra-heterotic group needs genetic improvement, and hybrids between inter-heterotic groups used to produce the superior hybrids. In an intra-group hybrid were observed high grain yield, SCA and heterosis than inter-group hybrids. The hybrid combination of an inter-group giving low yields but have high genetic distance, so genetic distance cannot as such predict the heterosis. However, the use of functional markers related to grain yield and its components would provide specific guidelines for combining parents correctly to increase breeding efficiency. Broadening the hybrid rice parental gene pools by

introducing germplasm from other sources and integrating them into heterotic groups are essential steps to further enhance the heterotic performance of hybrids in the state of Assam.

Mean comparison of heterotic groups based on genetic distances

Table (8) presents the mean performance of the hybrids in GD-based heterotic groups for the different traits. A comparative evaluation of heterotic group means for different traits could be suggested for improving specific characters. This comparison indicated that G2×G3 had better cluster means for most of the characters and therefore, G2×G3 needs consideration for selecting genotypes as parents in a hybridization programme. The hybrids of G1×G3 group could lead to early maturing hybrid development.

Seedling length (cm), Leaf number, Seedling establishment (%), days to panicle initiation, days to 50% flowering, days to maturity, culm length (cm), productive tillers, grains Panicle⁻¹, harvest index (%) and gel consistency (mm) recorded higher mean performances in the inter-group hybrids than in intra-group hybrids, providing more chances for segregation and recombination and thus, these traits concern priority in selecting promising parents for hybridization programme. For the remaining traits, mean performances of the intra-group hybrids were higher than that of the inter-group hybrids. The higher mean performance was evident for grain yield (g Plant⁻¹) in the intra-group hybrids than in the inter-group hybrids and was because of the adverse indirect effects of yield contributing traits on yield in the inter-group hybrids.

Correlation between heterosis and genetic distances

A perusal of Table (9) revealed that the parameters, namely, mean yield, HMP, HBP and HSP, were strongly correlated among themselves. Genetic distances showed no correlation with any of the mean yield and heterosis parameters. Grain yield per plant had positive and highly significant correlation ($p < 0.01$) with HMP (g Plant⁻¹), HMP (%), HBP (g Plant⁻¹), HBP (%) and HSP(%). HMP (g Plant⁻¹) had positive and significant association with HMP (%), HBP (g Plant⁻¹), HBP (%) and HSP (%). HMP (%) showed a positive and significant association with the HBP (g Plant⁻¹), HBP (%) and HSP (%). HBP (g Plant⁻¹) exhibited positive and significant association with HBP (%) and HSP (%). HBP (%) had a positive and significant association with the HSP (%). The correlation coefficient of genetic distance was negative and non-significant with grain yield per plant (-0.0067). Wang et al. 2015, also found the negative correlation of grain yield with the genetic distance. Singh 2015, found a weak association of grain yield with the genetic distance in maize. Many of the studies on rice found a negative correlation of genetic distance with the heterosis (Singh et al. 2011; Xu et al. 2002; Yu et al. 2011). The association of marker-based GD and hybrid performance was too small to be used for predicting hybrid breeding (Sant et al. 1999; Zhao et al. 2010). Tams et al. 2006 and Geleta et al. 2004, also confirmed the non-significant correlation of genetic distance with heterosis in maize. Thus, the molecular marker-based genetic distance may not always be reliable for rice hybrid breeding programme and prediction of heterosis. The information about functional markers related to yield heterosis might provide precise direction or guideline for combining parents definitely to increase breeding efficiency of rice. Zhang et al. 2010, suggested the prediction of heterosis based on yield-related functional genes. As the heterosis is measured mainly in terms of yield, the reliability of markers within genes (EST-SSRs) in heterosis prediction (Pavani et al. 2018). However, the recent results showed that the association and prediction could be enhanced when the parental groups are formed first by molecular markers, which may not predict the best hybrid combination, but it reveals a practical value of assigning existing and new hybrid rice germplasm into heterotic groups and increasing opportunities to

develop desirable hybrids from the best heterotic groups, which is consistent with a previous study in maize (Lanza et al. 1997). Pavani et al. 2018, suggested the use of molecular marker heterozygosity, combining ability, high mean performance for different traits, the morphological and molecular marker-based grouping of parental lines to identify heterotic patterns in rice.

Conclusions

In genetic distance based heterotic grouping, the intra-group hybrids were recorded a higher frequency of crosses, GYP, SCA, HMP, HBP and HSP value than those of inter-group hybrids. No correlation between heterosis and genetic distance can be attributable to the use of a subset of markers not linked to yield or concerned. Gene linked markers or high-density markers (genome-wide markers) or markers within yield genes should be used for genetic distance analysis and heterotic grouping based on genetic distances. Apart from the further study on the genetic aspects, it might be interesting to integrate epigenomics, metabolomics, proteomics, and systems biology approaches for gaining better understandings into the heterotic gene pools of rice. A careful selection of parents with desirable traits constellation contributing to high yield of the parents for better complementation should be emphasized along with a considerable genetic distance for augmentation of yield heterosis.

Acknowledgements

We are most thankful to the Department of Agricultural Biotechnology AAU-Jorhat for the SSR genotyping. We also acknowledge the contributions of Dr Debojit Sarma for their helpful comments and writing in the manuscript. Authors are very thankful to anonymous reviewers for their comments and suggestions.

1. Adhikari BN, Joshi BP, Shrestha J, Bhatta NR (2018) Genetic variability, heritability, genetic advance and correlation among yield and yield components of rice (*Oryza sativa* L.). J Agric Nat Resour 1(1):149-160
2. Anh T, ThiT, Khanh TD, Dat TD, Xuan TD (2018) Identification of phenotypic variation and genetic diversity in rice (*Oryza sativa* L.) mutants. Agriculture 8(2):30
3. Bhuiya MSU, Hossain SMA, Kabir SKG (1989) Nitrogen fertilization in rice cv. BR10 after green manuring. Bangladesh J Agril Sci 16(1):89-92
4. Borém A, Fritsche-Neto R, (2014) Biotechnology and plant breeding: applications and approaches for developing improved cultivars. Elsevier
5. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32(3):314
6. Cagampang GB, Perez CM, Juliano BO (1973) A gel consistency test for eating quality of rice. J Sci Food Agric 24(12):1589-1594
7. Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Am J Hum Genet 19:233-257
8. Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, Park WD, Ayres N, Cartinhour S, McCouch SR (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100(5):713-722
9. Choudhury DR, Singh N, Singh AK, Kumar S, Srinivasan K, Tyagi RK, Ahmad A, Singh NK, Singh R (2014) Analysis of genetic diversity and population structure of rice germplasm from the north-eastern region of India and development of a core germplasm set. PLOS one 9(11):1-12
10. Choukan R, Hossainzadeh A, Ghannadha MR, Warburton ML, Talei AR, Mohammadi SA (2006) Use of SSR data to determine relationships and potential heterotic groupings within medium to late maturing Iranian maize inbred lines. Field Crops Res 95:212-222
11. Dellaporta SL, Wood H, Hicks JB (1983) A plant DNA mini-preparation. Plant Mol Bio Rep 1(4):19-21
12. Ditta A, Zhou Z, Cai X, Shehzad M, Wang X, Okubazghi KW, Xu Y, Hou Y, Iqbal MS, Khan MKR, Wang K, Liu F (2018) Genome-wide mining and characterization of SSR markers for gene mapping and gene diversity in *Gossypium barbadense* L. and *Gossypium darwinii* G. watt accessions. Agron J 8:181
13. Ebaid RA, Ghanem SA (2000) Productivity of Giza 177 rice variety grown after different winter crops and fertilized with different nitrogen levels. Egypt J Agric Res 78:717-731
14. EL-Batal MA, Abd EL-Gawad MH, Abdo FA, EL-Set A, EL-Aziz A (2004) Uniconazole application as anti-lodging for rice plants fertilized with high nitrogen rate. Zagazig J Agric Res 31:473-490
15. Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Bio Evol 14(7):685-695

16. Geleta LF, Labuschagne MT, Viljoen CD (2004) Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. *Plant Breed* 123(5):467-473
17. Ghara AG, Nematzadeh G, Bagheri N, Ebrahimi A, Oladi M (2012) Evaluation of general and specific combining ability in parental lines of hybrid rice. *Int J Agric Res Rev* 2(4):455-460
18. Gour L, Koutu GK, Singh SK, Mishra DK (2017) Genetic potency identification of indigenous rice (*Oryza sativa* L.) lines using quality assessment and SSR marker analysis. *Int J Chem Stud* 5(4):97-101
19. Govt. of India [GOI], (2015) Ministry of Agriculture, (<https://databank.nedf.com/content/land-use-4>)
20. Gupta PK, Varshney RK, Sharma PC, Ramesh B (1999) Molecular markers and their application in wheat breeding. *Plant Breed* 118: 369-390
21. Hayes HK, Immer FR, Smith DC (1955) *Methods of plant breeding*. Mc. Grow Hill Book Co. Inc., New York, 555
22. He ZZ, Xie FM, Chen LY, Dela Paz MA (2012) Genetic diversity of tropical hybrid rice germplasm measured by molecular markers. *Rice Sci* 19:193-201
23. Hore DK (2005) Rice diversity collection, conservation and management in North-eastern India. *Genet Resour Crop Evol* 52:1129-40
24. Hussain T, Jilani G, Gaffer MA (1989) Influence of level and time of N application on growth and yield of rice. *Intl. Rice Res. Newsl.* 14(6):18
25. Islam MM, Hoque ME, Rabbi SMHA, Ali MS (2011) DNA fingerprinting and diversity Analysis of BRRI hybrid varieties and their corresponding parents. *Plant Tissue Culture and Biotechnology* 21(2):189-198
26. Islam MR, Haque MS, Bhuiya ZH (1990). Effect of nitrogen and sulphur fertilization on yield response and nitrogen and sulphur composition of rice. *Bangladesh J Agril Sci* 17(2):299-302
27. Jacob HJ, Lindpainter K, Lincoln SE, Kusumi K, Bunker RK, Mao Y, Ganten D, Dzau VJ, Lander ES (1991) Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. *Cell* 167:213-224
28. Kumar AS, Adilakshmi D (2016) Heterosis for yield, yield components and quality traits in rice hybrids (*Oryza sativa* L.). *Plant Arch* 16(1):237-242
29. Kumar I, Khush GS (1986) Genetics of amylose content in rice (*Oryza sativa* L.). *J Genet* 65:1-1
30. Kwon SJ, Ha WG, Hwang HG, Yang SJ, Choi HC, Moon HP, Ahn SN (2002) Relationship between heterosis and genetic divergence in 'Tongil'-type rice. *Plant Breed* 12:487-492
31. Lanza LLB, De Souza JCL, Ottoboni LMM, Vieira MLC, De Souza AP (1997) Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. *Theor Appl Genet* 94(8):1023-1030
32. Latha S, Sharma D, Sanghera GS (2013) Combining ability and heterosis for grain yield and its component traits in rice *Oryza sativa* L. *Not Sci Biol* 5(1):90-97
33. Li CD, Rossmagel BG, Scoles GJ (2000) The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. *Theor Appl Genet* 101:1259-1268

34. Liu K, Deng J, Lu J, Wang X, Lu B, Tian X, Zhang Y (2019) High nitrogen levels alleviate yield loss of super hybrid rice caused by high temperatures during the flowering stage. *Front Plant Sci* 10:357
35. Liu K, Muse SV (2005) Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128-2129
36. Lu Y, Yan J, Guimaraes CT, Taba S, Hao Z, Gao S, Chen S, Li J, Zhang S, Vivek BS, Magorokosho C, Mugo S, Makumbi D, Parentoni SN, Shah T, Rong T, Crouch JH, Xu Y (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. *Theor Appl Genet* 120(1):93-115
37. Mackill DJ, Coffman WR, Garrity DP (1996) Rainfed lowland rice improvement. Los Banos, The Philippines: International Rice Research Institute.
38. McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol Biol* 35:89-99
39. Melchinger AE, Gumber RK (1998) Overview of heterosis and heterotic groups in agronomic crops. *Concepts and breeding of heterosis in crop plants* 25:29-44
40. Ming H, Ya-hui WU, Xing-xing TAO, Yong-zhu L, Gui-li Y, Zhi-qiang C (2015) Genetic diversity of main inbred *Indica* rice varieties applied in Guangdong province as revealed by the molecular marker. *Rice Sci* 22(1):1-8
41. Mvuyekure SM, Sibiya J, Derera J, Nzungize J, Nkima G (2018) Assessment of genetic diversity of rice-based on SNP markers for selection of parents for sheath rot (*Sarocladiumoryzae*) resistance breeding. *S Afr J Plant Soil* 35(1):51-59
42. Pachauri V, Taneja N, Vikram P, Singh NK, Singh S (2013) Molecular and Morphological Characterization of Indian Farmers Rice Varieties (*Oryza sativa* L.). *Aust J Crop Sci* 7(7):923-32
43. Pavani M, Sundaram RM, Ramesha MS, Kishore K, Kemparaju KB (2018) Prediction of heterosis in rice based on the divergence of morphological and molecular markers. *J Genet* 97(5):1263-1279
44. Perrier X, Jacquemoud-Collet JP (2006) DARwin software: dissimilarity analysis and representation for windows Website <http://darwin.cirad.fr/darwin> [accessed 1 October 2019]
45. Pervaiz ZH, Rabbani MA, Shinwar ZK, Masood MS, Malik SA (2010) Assessment of genetic variability in rice (*Oryza sativa* L.) germplasm from Pakistan using RAPD markers. *Pak J Bot* 42(5):3369-3376
46. 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
47. Reif JC, Hallauer A, Melchinger A (2005) Heterosis and heterotic patterns in maize. *Maydica* 50:215-223
48. Rice knowledge bank. IRRI. (2002) <http://www.knowledgebank.irri.org>
49. Riday H, Brummer EC, Campbell TA, Luth D, Cazcarro PM (2003) Comparisons of genetic and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and subsp. *falcate*. *Euphytica* 131:37-45

50. Romay MC, Millard MJ, Glaubitz JC, Peiffer JA, Swarts KL, Casstevens TM (2013) Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol* 14(6): R55
51. Sahoo JP, Sharma V, Verma RK, Chetia SK, Baruah AR, Modi MK, Yadav VK (2019) Linkage analysis for drought tolerance in Kharif rice of Assam using microsatellite markers. *IJTK* 18(2):371-375
52. Sahu R, Singh SK, Singh DK, Vennela PR, Yerva SR, Kumar D, Singh M, Rathan ND (2017) Heterosis studies for yield & yield traits in rice (*Oryza Sativa* L.) under rainfed condition. *IJEAB* 10(1):1-10
53. Salem, A.K.M. (2006). Effect of nitrogen levels, plant spacing and time of farmyard manure application on the productivity of rice. *J. Appl. Sci. Res.*, 2: 980-987
54. Sant VJ, Patankar AG, Sarode ND, Mhase LB, Sainani MN, Deshmukh RB, Ranjekar PK, Gupta VS (1999) Potential of DNA markers in detecting divergence and in analysing heterosis in Indian elite chickpea cultivars. *Theor Appl Genet* 98:1217-1225
55. Seetharam K, Thirumeni S, Paramasivam K (2009) Estimation of genetic diversity in rice (*Oryza sativa* L.) genotypes using SSR markers and morphological characters. *Afr J Biotechnol* 8(10):2050-2059
56. Shamim MZ, Manzar H, Sharma VK, Kumar P (2015) Microsatellite marker-based characterization and divergence analysis among rice varieties. *IJBT* 15:182-189.
57. Sharma V, Saini DK, Kumar A, Kesh H, Kaushik P (2020) Breeding for Biofortification Traits in Rice: Means to Eradicate Hidden Hunger. In *Agronomy*. IntechOpen.
58. Singh N, Choudhury DR, Tiwari G, Singh AK, Kumar S, Srinivasan K, Tyagi RK, Sharma AD, Singh NK, Singh R (2016). Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. *BMC Genet* 17:127
59. Singh PR (2015) Genetic distance, heterosis and combining ability studies in maize for predicting F₁ hybrid performance. *Sabrao J Breed Genet* 47(1):21-8
60. Singh RK, Sharma RK, Singh AK, Singh VP, Singh NK, Tiwari SP, Mohapatra T (2004) Suitability of mapped sequence-tagged microsatellite site markers for establishing distinctness, uniformity and stability in aromatic rice. *Euphytica* 135(2):135-143
61. Singh SK, Singh CM, Lal GM (2011) Assessment of genetic variability for yield and its component characters in rice (*Oryza sativa* L.). *Res Plant Biol* 1(4):73-76
62. Sivaranjani AKP, Pandey MK, Sudharshan I, Kumar GR, Madhav MS, Sundaram RM (2010) Assessment of genetic diversity among basmati and non-basmati aromatic rice of India using SSR markers. *Curr Sci* 99:221-226
63. Tams SH, Bauer E, Oettler G, Melchinger AE, Schon CC (2006) Prospects for hybrid breeding in winter triticale: II. Relationship between parental genetic distance and specific combining ability. *Plant Breed* 125:331-336
64. Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovtiesich L, et al (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor Appl Genet* 100:697-712
65. Thorat BS, Kunkerkar RL, Bagkar TA (2017) Studies on heterosis for yield and its contributing traits in hybrid rice (*Oryza sativa* L.). *Int J Chem Stud* 5:7-12

66. Travis AJ, Norton GJ, Datta S, Sarma R, Dasgupta T, Savio FL, Macaulay M, Hedley PE, McNally KL, Sumon MH, Islam MR (2015) Assessing the genetic diversity of rice originating from Bangladesh, Assam and West Bengal. *Rice* 8(1):35
67. Turner JM (1953) A study of heterosis in upland cotton II Combining ability and inbreeding effects. *Agron J* 43:487-490
68. Vanniarajan C, Vinod KK, Pereira A (2012) Molecular evaluation of genetic diversity and association studies in rice (*Oryza sativa* L.). *J Genet* 91:9-19
69. Verma H, Borah JL, Sarma RN (2019) Variability assessment for root and drought tolerance traits and genetic diversity analysis of rice germplasm using SSR markers. *Sci Rep* 9(1):1-19
70. Virk PS, Newbury JH, Bryan GJ, Jackson MT, Ford-Lloyd BV (2000) Are mapped or anonymous markers more useful for assessing genetic diversity? *Theor Appl Genet* 100:607-613
71. Wang K, Qiu F, Dela Paz WLMA, Xie F (2015) Heterotic groups of tropical *indica* rice germplasm. *Theor Appl Genet* 128:421-430.
72. Wang Y, Lu J, Ren T, Hussain S, Guo C, Wang S, Cong R, Li X (2017) Effects of nitrogen and tiller type on grain yield and physiological responses in rice. *AoB Plants* 9: plx012; doi:10.1093/aobpla/plx012
73. Watanabe KN, Ohsawa R, Obara M, Yanagihara S, Aung PP, Fukuta Y (2016) Genetic variation of rice (*Oryza sativa* L.) germplasm in Myanmar based on genomic compositions of DNA markers. *Breed Sci* 16033
74. Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388-396
75. Westman AL, Kresovich S (1997) Use of molecular marker techniques for the description of plant genetic variation. *Biotechnology in agriculture series* 9-48
76. Xie C, Warburton M, Li M, Li X, Xiao M, Hao Z, Zhao Q, Zhang S (2008) Retracted article: an analysis of population structure and linkage disequilibrium using multi locus data in 187 maize inbred lines. *Mol Breeding* 21(4):407-418
77. Xie F, He Z, Esguerra MQ, Qiu F, Ramanathan V (2014) Determination of heterotic groups for tropical Indica hybrid rice germplasm. *Theor Appl Genet* 127(2):407-417
78. Xu W, Virmani SS, Hernandez JE, Sebastian LS, Redoña ED, Li Z (2002) Genetic diversity in the parental lines and heterosis of the tropical rice hybrids. *Euphytica* 127:139-148
79. Yadav S, Anuradha G, Kumar RR, Vemireddy LR, Sudhakar R, Donempudi K, Venkata D, Jabeen F, Narasimhan YK, Marathi B, Siddiq EA (2015) Identification of QTLs and possible candidate genes conferring sheath blight resistance in rice (*Oryza sativa* L.). *Springerplus* 4(1):175
80. Yingheng W, Qihua CAI, HongguangX, Fangxi WU, Ling L, Wei H, Liping C, Hua'an X, Jianfu Z (2018) Determination of heterotic groups and heterosis analysis of yield performance in indica Rice. *Rice Sci* 25(5):261-269
81. Yu Y, Yuan DJ, Liang SG, Li XM, Wang XQ, Lin ZX, Zhang XL (2011) Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC₁ population between *Gossypium hirsutum* and *G. barbadense*. *BMC Genomics* 12(1):15
82. Zeng L (2004) Response and correlated response to salt tolerance selection in rice by yield parameters. *Cereal Res Commun* 32:477-484

83. Zhang T, Nixian L, Jiang KF, Yang QH, Yang L, Wan XQ (2010) Correlation between heterosis and genetic distance based on molecular markers of functional genes in rice. *Rice Sci* 17:1-7
84. Zhao FJ, Ma JF, Meharg AA, McGrath SP (2009) Arsenic uptake and metabolism in plants. *New Phytol* 181:777-794
85. Zhao K, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, Tyagi W, Ali ML, Tung CW, Reynolds A, Bustamante CD (2010) Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. *PloS one* 5(5)

Figure 1 Hierarchical horizontal clustering of 60 *Ahu rice* genotypes of Assam using Unweighted Neighbour-joining (UNJ) method based on chord distances estimated from SSR marker data

Figure 2: Hierarchical axial clustering of 11 *Ahu rice* genotypes using Unweighted Neighbour-joining (UNJ) method based on genetic distances estimated from SSR markers

Figure 3: Frequency distribution of genetic distances of the 60 rice genotypes and the selected 11 parents for the diallel crosses based on 53 SSR markers

Fig 1: Gel pictures of some of the polymorphic markers showing the amplified products