Oryza Vol. 54 No. 2, 2017 (174-185)
DOI 10.5958/2249-5266.2017.00023.6

Assessment of polymorphism at molecular level, association studies, multivariate analysis and genetic diversity among recombinant inbred lines of rice (*Oryza sativa* L.)

S Mohanty, N Mohanty, S Sulakshana, SK Pradhan, SK Dash and L Behera*

ICAR- National Rice Research Institute, Cuttack-753006, Odisha, India

****Corresponding author e-mail: l_behera@yahoo.com***

Received : 01 June 2017 Accepted : 27 September 2017 Published : 28 September 2017

ABSTRACT

Five hundred twenty nine microsatellite markers were used to assess polymorphism at molecular level between six rice cultivars, PDKV Shriram, Heera, AC38562, Pimpudibasa, Reeta and WAB56-50 having wide variation in yield and related traits. Ninety four (17.76%) microsatellite markers showed polymorphism between these six cultivars. Maximum polymorphism was detected between Reeta and WAB56-50 (34.02%), followed by AC38562 and Pimpudibasa (31.76%), and PDKV Shriram and Heera (27.22%). The recombinant inbred line (RIL) mapping population developed from Reeta and WAB56-50 were used for assessing genetic variability based on the yield and its component traits. The analysis revealed the significant difference among the RILs for all the yield traits. Correlation and path analysis revealed the strong association of grain yield with traits like grain number, thousand grain weight, panicle length and total spikelets per panicle. Further, principal component analysis indicated 58.12% of the total variation explained by first four principal components. Therefore, these RILs could be used for mapping of QTLs associated with yield and its component traits. The polymorphic markers identified in the present study would be useful for mapping QTLs associated with yield and its component traits.

Key words: *Genetic variability, microsatellite, principal component analysis, recombinant inbred lines*

INTRODUCTION

Rice (*Oryza sativa* L.) is one of the most important cereal crops in the world and serves as the major food for almost half of the global population. It is cultivated in 150 million ha annually which is 11 % of the world's cultivated land. More than 90% of rice is produced and consumed in Asia. India alone possesses nearly one third of rice growing area covering 44.5 million hectares. According to RMM (Rice Market Monitor), the world rice production is 758.8 million tons (503.6 million tons, milled basis) in 2017 which is 0.8 percent (5.8 million tons) higher than that of 2016.With ever increasing population growth, limited land and water resources, urbanization, post-harvest losses, global warming and continued emergence of new pests and diseases, there is a growing demand to increase the productivity of rice crop. Enhancing grain yield is a key solution to support ever increasing population. Rice has a highly diverse genetic make-up, including inferior and superior alleles. Pyramiding favorable alleles of quantitative trait loci (QTLs) for yield traits would make a significant contribution to breeding high-yielding rice varieties.

Grain yield and its related traits are the most important quantitative traits that control the productivity of rice. Inheritance pattern of quantitative traits is not only very complex to study but also assumed to be severely affected by environmental factors. Most of the traits of interest are complex and are result of the interaction of a number of component traits (Sarawgi et al., 1997). Hence, selection based on the phenotype would be very difficult and often unreliable. Manipulation of such traits in breeding program is very difficult and often unreliable because they are controlled by several inter-component traits.

Grain yield is dependent on many yield component traits as well as on the environmental factors. Genetic variability of yield contributing traits and interrelationship among them are necessary for a successful breeding program. Knowledge of heritability is essential for selection based improvement. Before placing strong emphasis on breeding for yield improvement trait, the knowledge on the association between yield and yield attributes will immense help to the breeder in the improvement of yield. The correlation coefficient may also help to identify characters that have little or no importance in the selection program. The existence of correlation may be attributed to the presence of linkage or pleiotropic effect of genes or physiological and development relationship or environmental effect or in combination of all (Oad et al., 2002). Path coefficient analysis proposed by Wright (1921) help the partition of the total correlation into direct and indirect effects of various causes.

Molecular marker technology provides powerful tool for assessment of genetic diversity among cultivars, identification of cultivars, and thus add to management and protection of plant genetic resources (Virk et al., 2000). Furthermore, molecular marker technology is highly reliable, rapid (2-3 days compared with many seasons by field technique), requires less capital expenditure and can be used for transfer more than one gene/QTL into desired background simultaneously. Of the wide array of DNA markers available, microsatellite markers are considered to be appropriate for assessment of genetic diversity and variety identification because of their ability to detect large numbers of discrete alleles repeatedly, accurately and efficiently (Smith et al., 1996; Anandan et al., 2016; Pradhan et al., 2016; Pandit et al., 2017). DNA fingerprinting/ profiling is used as versatile tool for investigating various aspects of plant genomes including characterization of genetic variability, genome fingerprinting, genome mapping, gene localization, analysis of genome evolution, population genetics, taxonomy and plant breeding. Molecular data provides a basis for better management and conservation of the collection and could be used as reference for its enhanced use in breeding programs.

In the present study, 529 microsatellite markers were used to assess polymorphism between six rice cultivars (PDKV Shriram, Heera, AC38562, Pimpudibasa, Reeta and WAB56-50) differing widely in yield traits and to estimate the genetic variability in a RIL mapping population developed from these two highly diverse cultivars, Reeta and WAB56-50.

MATERIALS AND METHODS

Plant materials

The experimental material of the present study consists of six rice cultivars, namely PDKV Shriram (high grain number), Heera (low grain number), AC38562(high 1000-grain weight), Pimpudibasa (low 1000-grain weight), Reeta (high *per se* yield) , WAB56-50 (low *per se* yield) and 190 F_9 recombinant inbred lines developed from the cross between Reeta and WAB56-50. A careful cross was made between Reeta (as female parent) and WAB 56-50 (as male parent) and F_1 seeds were collected. F_1 plants were grown in the pots and hybridity (heterozygocity) was checked using microsatellite markers. A single F_1 plant was selfed to produce F_2 plants. The F_2 plants were advanced to F_9 recombinant inbred lines (RILs) following single seed descent during 2009-2015. Finally, 190 recombinant inbred lines (RILs) were used for our study.

Assesment of DNA polymorphism between rice cultivars DNA isolation and PCR amplification

Twenty seeds from parents and each RIL were germinated in Petridis. Two hundred to three hundred milligrams of leaves from 10-15 young seedlings were collected and genomic DNA was isolated following Cetyl Trimethyl Ammonium Bromide (CTAB) method (Murray et al., 1980). A set of 529 microsatellite markers uniformly distributed over 12 rice chromosomes were used to assess polymorphism between six rice cultivars. The PCR amplification was carried out in a 20µl reaction mixture volume containing 40ng of genomic DNA, 1X PCR buffer {75 mM Tris-HCl (pH 9.0), 50mM KCl, 20 mM $(NH_4)_2SO_4$}, 200µM dNTP mix (MBI Fermantas, Lithuania, USA), 4 picomole of each of forward and reverse primers, 2 mM of $MgCl_2$ and 1U of Taq (Thermus aquaticus) DNA polymerase (Biotools, Spain). The PCR was performed in a thermal cycler (Applied Biosystems, USA) as per following cycling parameters: initial denaturation at 94^0C for 3 min followed by 36 cycles of denaturation at 94^0C for 1min, annealing at 55-67^0C (depending upon primer) for 1 min and extension at

72^0C for 1.5 min and final extension at 72^0C for 5 min. The amplified products were separated on 2.5% agarose gels using 1X TBE buffer and stained with ethidium bromide (0.5µg/ml). The gels were visualized under UV radiation and photographed using a gel documentation system (Fluor Chem™ 5500, Alpha Innotech, USA) to detect polymorphism. The size of amplified bands was determined based on the migration relative to molecular weight size markers (50 bp DNA ladder, MBI Fermentas, Lithuania) using Alphaease software (Alpha Innotech, USA).

Evaluation of mapping population (recombinant inbred lines) for yield and its component traits

The RIL mapping population along with parents, Reeta and WAB56-50 were evaluated for yield and its component traits at experimental fields of National Rice Research Institute (NRRI), Cuttack during *Kharif*, 2015 following Alpha Lattice design. Twenty five days old seedlings of both parents and individual RIL were transplanted in the experimental plots. Each entry was planted with 20 single plants per row at spacing of 20 cm between rows and 15 cm between plants. Gap filling was done within a week in order to maintain uniform plant population. Fertilizer dose of 80 kg N, 40 kg P_2O_5, and 40 kg K_2O was applied. Entire dose of P_2O_5 and K_2O along with half dose of N was applied as basal dose at the time of final field preparation. Remaining amount of nitrogen was splitted in two equals and were applied at the time of tillering and grain filling stages. The standard agronomic practices were adopted for normal crop growth. Five plants were selected from middle of the row for data collection. The data on 14 traits recorded were days to 50% flowering (DFF), plant height (PH), number of tillers/ plant (TN), number of panicles/ plant (PN), panicle length (PL), number of fertile grains per panicle (GN), number of chaff per panicle (CN), number of spikelets per panicle (SN), spikelet fertility percentage (SFP), number of primary branching/ panicle (PB), number of secondary branching/panicle (SB), panicle weight (PW), 1000-grain weight (GW) and *per se* yield (YLD).

Statistical analysis

The amplified bands were scored for each microsatellite loci based on the presence or absence of bands, generating a binary data matrix of 1 and 0 for each marker system. Total numbers of bands within each lane and polymorphic bands were noted. Each band was scored as an allele. These binary data matrix was then utilized to generate genetic similarity data among RILs.

The percentage of polymorphism was calculated using the formula:

% of polymorphism = (No. of polymorphic loci/ Total No. loci used)X 100

The statistical analysis on the mean values of five randomly selected plants from each of the two replications for RILs and parents was carried out on individual traits. The data of mean value for all the traits were analyzed for their variance following simple lattice design outlined by Meier (1954). Analysis was done using WINDOSTAT statistical package. The significance was tested by referring to the table given by Fisher (1936). Range, phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), heritability and genetic advance were calculated comparing means of RILs. The phenotypic correlations were estimated from variance and covariance components practiced by Sarawgi et al. (1997). Partitioning of correlation coefficients of traits into direct and indirect effects was carried out using the procedure practiced by Surek and Beser (2003). The degree of association of component traits with the dependent trait like grain yield and the correlation coefficients were computed using the formula given by Al-Jobourie et al. (1958). Path coefficient analysis was carried out using phenotypic coefficient to ascertain the direct and indirect effect of the yield components on yield as suggested by Wright (1921) and Dewey and Lu (1959). The path coefficient analyses were conducted as described by Williams et al. (1990) for the data. Principal component analysis was done with PAST3.0 statistical package. RILs used in the present study were statistically analyzed for similarities in their quantitative traits using cluster observations analysis by Ward method (Ward, 1963).

RESULTS AND DISCUSSION

Assessment of DNA polymorphism between rice cultivars

The accurate assessment of genetic diversity is important not only for crop improvement but also for

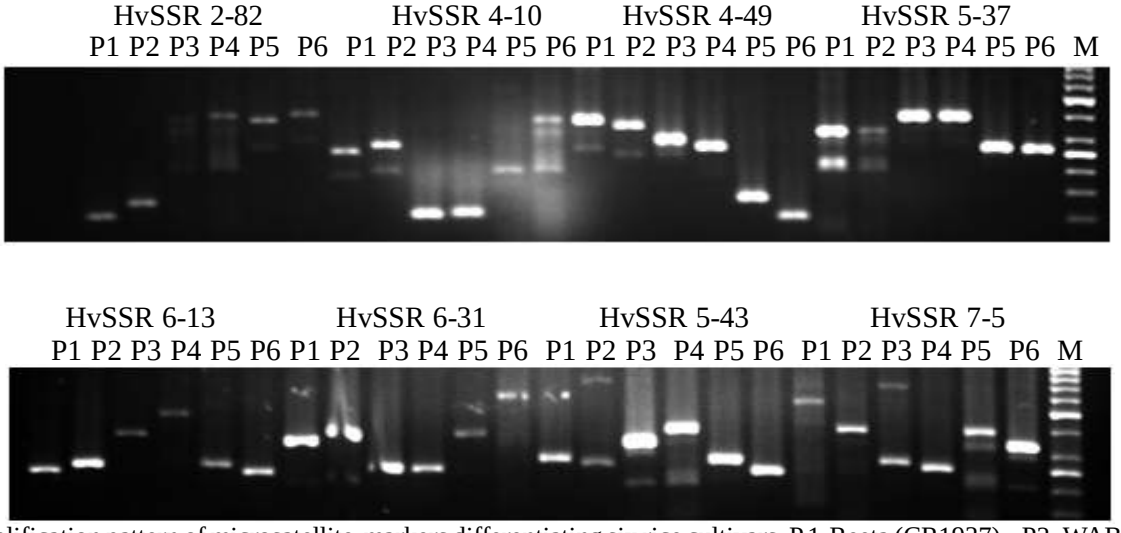

Fig. 1. Amplification pattern of microsatellite markers differentiating six rice cultivars. P 1-Reeta (CR1937), P2- WAB56-50 , P3- PDKV Shriram, P4- Heera, P5-AC 35862, P6- Pimpudibasa, M- 50bp DNA Ladder .

efficient management and protection of the germplasm. Further, this will maximize the probability of transgressive segregation and accumulation of positive alleles from different cultivars. Microsatellites are considered to be appropriate for assessment of genetic diversity, fingerprinting for variety identification and assessment of seed purity (Behera et al., 2012). Five hundred twenty nine microsatellite markers were used to asses polymorphism between six rice cultivars. Ninety four (17.76%) microsatellite markers showed polymorphism between all the six cultivars (*i.e.*, PDKV Shriram, Heera, AC38562, Pimpudibasa, IET1996 and WAB50-56). One hundred forty four (27.22%) were found to be polymorphic between PDKV Shriram and Heera while 168 (30.88%) microsatellite markers differentiated AC38562 and Pimpuidbasa. One hundred eighty (34.02%) microsatellite markers were polymorphic between Reeta and WAB50-56 (Fig. 1), indicating that Reeta and WAB50-56 are genetically more diverse as compared to other pairs. Reeta is a late maturing, semi-dwarf and *indica* variety developed through hybridization involving Savitri/ IR44. WAB 56-50 is a well known upland NERICA rice cultivar developed from the hybridization between IDSA6/ IAC

Table1. Microsatellite markers used for assessment of polymorphism between rice Cultivars.

Chrom#	Reeta and WAB (56-50)			PDK Sriram and Heera			AC 38562 and Pimpudibasa		
	Microsatellite marker#	Polymorphic microsatellite Marker#	% polymorphism	Microsatellite marker#	Polymorphic microsatellite Marker#	% polymorphism	Microsatellite marker#	Polymorphic microsatellite Marker#	% polymorphism
1	67	25	37.31	67	17	25.37	67	19	28.35
2	62	26	41.93	62	19	30.64	62	17	27.41
3	56	18	32.14	56	15	26.78	56	24	42.85
4	44	14	31.81	44	9	20.45	44	14	31.81
5	40	12	30.00	40	11	27.50	40	12	30.00
6	47	13	27.65	47	18	38.29	47	15	31.91
7	31	11	35.48	31	7	22.58	31	13	27.66
8	64	14	21.87	64	14	21.87	64	17	26.56
9	31	12	38.70	31	12	38.70	31	12	38.70
10	27	8	29.62	27	11	40.74	27	8	29.63
11	30	9	30.00	30	10	33.33	30	8	26.67
12	31	8	25.80	31	9	29.03	31	9	29.03
Total	529	180	382.31	529	144	355.28	529	168	370.58
Average	44.08	15.00	34.02	44.08	12	27.22	44.08	14	30.88

164. This variety has short height, short growth cycle and high tillering capacity having genetic back ground of *glaberima*. Hence, they showed highest percentage (34.02%) of polymorphism. AC3862 is a tropical *japonica* variety while Pimpudibasa is a aromatic variety and showed 30.88% polymorphism. PDKV Shriram and Heera are high grain and low grain *indica* rice varieties, respectively and showed comparatively lower percentage (27.22%) of polymorphism (Table 1). Among the 12 chromosomes analyzed, the chromosome 2 showed highest polymorphism between Reeta and WAB 50-56 (41.93%) followed by chromosome 9 and 1. Similarly, chromosome 10 showed highest polymorphism between PDKV Shriram and Heera (40.74%) followed by chromosome number 9 and 6. AC38562 and Pimpudibasa showed highest polymorphism on chromosome 3 (42.85%) followed by chromosome number 9 and 6. In our study, the amplified product size ranged from 60 to 1000 bp in all the six genotypes. Chromosome 7 (25.80%) revealed maximum polymorphism between cultivars as compared to other chromosomes. Mohanty et al. (2017) reported 13.18% of polymorphism between BPH resistant parent Salkathi and susceptible parent TN1. Higher polymorphism was detected when *indica* and *japonica* parents were used for mapping of genes, *bph2* (28% between ASD7 and C418; Sun et al., 2006), *Bph9* (34% between Kaharamana and 02428; Su et al., 2006), *Bph12* (37.1% between B14 and TN1; Yang et al., 2002) and *Bph17* (32.5% between Rathu Heenati and 02428; Sun et al., 2005). Based on the polymorphism, a RIL mapping population was developed between Reeta and WAB50-56 which was used for assessing genetic diversity for yield and its component traits.

Table 2. Estimates of genetic parameters for grain yield and its component traits in RILs.

Characters	Mean	Range	GCV	PCV	H^2	GAM
DFF	125.8	93.0-149.0	8.17	8.2	98.0	16.7
PH	117.0	75.3-176.6	15.9	16.2	96.0	32.2
TN	7.6	2.6	17.7	21.2	70.0	30.7
PN	7.0	4.1-11.7	15.7	20.01	62.0	25.6
PL	24.7	16.5-32.9	11.5	13.25	75.0	20.6
GN	150.5	54.6-289.6	25.7	27.13	90.0	50.1
CN	26.9	6.1-104.5	58.5	60.7	93.0	116.4
SN	177.4	74.7-321.2	23.3	24.6	90.0	45.6
SFP	84.6	56.5-96.9	9.3	9.7	92.0	18.4
PB	12.1	6.7-55.5	33.7	36.3	86.0	64.5
SB	27.8	10.0-69.4	28.7	29.7	93.0	57.1
PW	2.9	1.2-5.8	28.4	30.8	85.0	54.0
TGW	28.7	11.6-46.2	27.1	27.6	96.0	54.8
YLD	17.9	5.4-32.1	27.0	32.5	69.0	46.4

DFF- Days to 50 % flowering, PH - Plant height (cm), TN - Tiller number/plant, PN - Panicle number/plant, GN-Fertile grain number/panicle, CN -Chaff number/panicle, SN- Spikelet numbers/panicle, SFP - Spikelet fertility percentage, PL-Panicle length (cm), PB - Primary branching, SB-Secondary branching, PW- Panicle weight (gm), TGW- 1000 grain weight (gm), YLD- Per se yield (gm)

Assessment of genetic diversity among RILs

The analysis of variance revealed highly significant difference among the RILs for all the 14 yield traits indicating a large amount of genetic variability is present in RIL population for effective selection and QTL mapping. The genotypic correlations were slightly higher in magnitude than corresponding correlations at phenotypic level (Table 2). Khedikar et al. (2004) and Kumar et al. (2013) also reported higher estimates of genotypic correlations than the corresponding phenotypic correlations between yield and its components in the rice RILs. Chaff number is the only character which showed strong negative associations with grain yield at genotypic and phenotypic level. The highest estimate of PCV and GCV were observed for chaff number (60.7, 58.5), primary branching (36.3, 33.7), panicle weight (30.8, 28.4) and grain yield per plant (32.5, 27), while the lowest in days to 50 per cent flowering (8.2, 8.17). Similarly, Bhadru et al. (2012) reported high PCV and GCV for number of grains per panicle, fertility percentage and grain yield per plant in rice. The estimate of heritability were high for days to 50% flowering (98%), plant height (96%), grain number (90%), chaff number (93%), thousand grain weight (96%) and spikelet fertility percentage (92%) due to genetic causes rather only by environmental effects. High heritability does not always indicate high genetic gain; heritability coupled with high genetic advance should be used in predicting the ultimate effect for selecting superior varieties. High heritability along with high genetic advance as per cent of mean was recorded for chaff number, primary branching, panicle weight, thousand grain weight and total number of grains per panicle indicated that the less influence of environmental effect in the inheritance of these traits. High heritability coupled with low genetic advance as per cent mean were observed in days to 50% flowering. Similar results were also reported by Gangashetty et al. (2013).

Table 3. Estimates of phenotypic (P) and genotypic (G) correlation coefficients between different yield and its component traits in recombinant inbred lines.

	DFF	PH	TN	PN	PL	GN	CN	SN	SFP	PB	SB	PW	TGW	YLD
Characters	1.000	-0.035	-0.049	0.028	0.006	0.004	0.076	-0.025	0.092 *	0.016	0.060	0.011	0.1266**	0.028
G	1.000	-0.028	-0.053	0.019	-0.001	-0.016	-0.066	-0.040	0.087	0.032	0.066	0.026	0.140	0.036
PH -P		1.000	0.017	0.026	0.320***	-0.065	0.179**	0.143***	-0.052	0.112 **	0.172***	0.1292 **	0.143***	0.196*
G		1.000	0.027	0.018	0.383***	0.178**	-0.072	0.137	0.108	-0.074	0.175*	0.121	0.126	0.255**
TN-P			1.000	0.501***	0.014	0.080	-0.058	0.053	0.056	-0.143***	-0.004	-0.042	0.007	0.167*
G			1.000	0.506***	0.055	0.114	-0.094	0.070	0.099	-0.174*	0.007	-0.081	0.025	0.170**
PN-P				1.000	0.054	0.1248 * *	- 0.0903 *	0.082*	0.128**	-0.152***	-0.023	0.1029 *	0.032	0.209*
G				1.000	0.068	0.175**	-0.086	0.129	0.147*	-0.165*	-0.019	0.133	0.021	0.242**
PL-P					1.000	0.186***	0.010	0.177***	0.049	-0.026	0.154 * **	0.1370***	0.133 **	0.160*
G					1.000	0.259**	-0.015	0.234**	0.093	-0.044	0.111	0.141	0.177*	0.218**
GN-P						1.000	-0.016	0.927***	0.430***	0.093*	0.220 * **	0.3974***	0.192 ***	0.273*
G						1.000	0.000	0.923***	0.408***	0.081	0.229**	0.420***	0.164*	0.307***
CN-P							1.000	0.359***	-0.876***	0.040	0.088 *	-0.065	0.070	-0.013
G							1.000	0.384***	-0.884***	0.047	0.098	-0.031	0.030	-0.028
SN-P								1.000	0.074	0.102*	0.238 * **	0.3467 * **	0.205 ***	0.250*
G								1.000	0.037	0.093	0.249**	0.376***	0.163*	0.273**
SFP-P									1.000	0.003	0.013	0.2232***	0.045	0.112*
G									1.000	-0.011	0.002	0.198**	0.069	0.153*
PB-P										1.000	0.482***	0.1509***	-0.122**	-0.070
G										1.000	0.512***	0.124	-0.131	-0.103
SB-P											1.000	0.3069***	0.035	0.070
G											1.000	0.297***	0.037	0.091
PW-P												1.000	0.085 *	0.168*
G												1.000	0.075	0.226**
TGW-P													1.000	0.199*
G													1.000	0.204**
YLD-P														1.000
G														1.000

*Significant 5% level, ** Significant 1% level, ***Significant 0.1% level
DFF - Days to 50 % flowering, PH - Plant Height (cm), TN – Tiller Number, PN – Panicle Number, PL – Panicle Length, GN – Grain Number, CN –Chap Numbers, SN– Spikelet Numbers, SFP – Spikelet Fertility Percentage, PB – Primary Branching, SB- Secondary Branching, PW- Panicle Weight, TGW- 1000 Grain Weight, YLD- *Perse* Yield

Character association and path analysis

The estimates of phenotypic and genotypic correlation coefficient are presented in Table 3. Grain yield displayed positive and highly significant association with all the yield component traits such as plant height (0.20), tiller number (0.17), panicle number (0.21), panicle length (0.16), grain number (0.27), spikelet number (0.25), spikelet fertility percentage (0.11), panicle weight (0.17) and thousand grain weight (0.20) at phenotypic level. Previous papers have mentioned rice grain yield as a function of maximum tiller density, number of filled grains per panicle (Samonte et al., 1998), and panicle length (Iftekharuddaula et al., 2002). The phenotypic correlation between grain yield with days to 50% flowering (0.03) and secondary branching (0.07) was positive but not significant at 0.05 level while the phenotypic correlation of grain yield with chaff number (-0.01) and primary branching (-0.07) was negative and not significant. At genotypic level grain yield displayed positive but no significant correlation with days to 50% flowering (0.0328), plant height (0.2284), tiller number (0.2398), panicle number(0.3276), panicle length (0.2069), grain number(0.3472), spikelet number (0.3218), spikelet fertility percentage (0.1388), secondary branching (0.0893), panicle weight (0.1987) and thousand grain weight (0.2371). The genotypic correlation of grain yield with chap number (-0.0055) and primary branching (-0.1081) was negative and not significant. The genotypic correlation coefficient was found to be higher than phenotypic correlation coefficient except for days to maturity indicating a strong inherent association for grain yield per plant and other traits also. A very strong positive correlation of grain yield per plant at genotypic and phenotypic level was observed with days to 50% flowering, plant height, tiller number, panicle length, grain number, spikelet fertility percentage, panicle weight and thousand grain weight. The positive association of grain yield with the traits mentioned above has been observed by various workers (Besufikad et al., 2016; Kayvan et al., 2007). The association studied indicating grain yield of rice can be improved by selecting lines having higher performances for these traits. Path co-efficient analysis measures the direct and indirect effect for one variable upon another and permits the separation of the correlation co-efficient into components of direct and indirect effect (Dewey and Lu, 1959).

Path coefficient analysis of yield and its components revealed that spikelet number had strong direct positive effect of 2.96 on grain yield. However spikelet number had indirect positive effect through plant height (0.4), tiller number (-0.06), panicle number (0.2), panicle length (0.5), grain number (2.7), chaff number (1.06), spikelet fertility percentage (0.2), primary branching (0.3), secondary branching (0.7), panicle weight (1.02) and thousand grain weight (0.6). Days to 50% flowering (-0.07) had negative indirect effect on grain yield. Similar results were reported by Ibrahim et al. (1990) and Iftekharuddaula et al. (2002). Similarly thousand grain weight had strong and direct positive effect of 0.13 with grain yield. It had a positive indirect effect through chaff number (0.0) and secondary branching (0.0003). Days to 50% flowering (-0.001), plant height (-0.0002), tiller number (-0.0005), panicle number (-0.0004), panicle length (-0.0006), grain number (-0.0007), spikelet number (-0.0006), spikelet fertility percentage (-0.0006), primary branching (-0.0007) and panicle weight (-0.0008) had negative indirect effect on grain yield. Following to these panicle number (0.12) and plant height (0.11) had strong direct positive effect on gain yield. Panicle number showed positive indirect effect via plant height (0.01), tiller number (0.20), panicle length (0.03), grain number (0.04), spikelet number (0.04), spikelet fertility percentage (0.02), primary branching (0.04), panicle weight (0.01) and thousand grain weight (0.03). Days to 50% flowering (-0.003), chaff number (-0.003), and secondary branching (-0.06) had negative indirect effect on grain yield. Similarly plant height had positive indirect effect via days to 50% flowering (0.0002), tiller number (0.005), panicle number (0.007), panicle length (0.06), grain number (0.02), chap number (0.004), spikelet number (0.02), spikelet fertility percentage (0.01), primary branching (0.03), secondary branching (0.01), panicle weight (0.03), thousand grain weight (0.009). Negligible direct effects of (0.08), (0.06), (0.05) and (0.02) were recorded for tiller number, panicle weight, panicle length and days to 50% flowering, respectively. Grain number (-2.5), spikelet fertility percentage (-0.19), chaff number (-1.3) and primary branching (-0.04) showed negative direct effect on grain yield per plant. While grain number had a positive indirect effect through plant height (0.14), tiller number (0.1), panicle number (0.13), panicle length (0.3), chaff number (0.008), spikelet number (0.6), spikelet fertility

Table 4. Direct and indirect effect of component traits on grain yields in RILs

Characters	DFF	PH	TN	PN	PL	GN	CN	SN	SFP	PB	SB	PW	TGW	YLD
DFF	0.022	-0.001	-0.001	0.001	0.000	0.000	-0.002	-0.001	0.002	0.000	0.001	0.000	0.003	0.028
PH	-0.004	0.112	0.002	0.003	0.036	0.020	-0.007	0.016	0.013	-0.006	0.019	0.015	0.016	0.196
TN	-0.004	0.001	0.080	0.040	0.001	0.006	-0.005	0.004	0.005	-0.012	0.000	-0.003	0.001	0.168
PN	0.004	0.003	0.062	0.124	0.007	0.016	-0.011	0.010	0.016	-0.019	0.003	0.013	0.004	0.210
PL	0.000	0.017	0.001	0.003	0.053	0.010	0.001	0.009	-0.003	0.001	0.008	0.007	0.007	0.160
GN	-0.009	-0.449	-0.201	-0.312	-0.467	-2.503	0.039	-2.321	-1.077	-0.233	-0.551	-0.994	-0.481	0.274
CN	0.095	0.082	0.073	0.114	-0.012	0.020	-1.261	-0.453	1.104	-0.050	-0.111	0.082	-0.088	-0.013
SN	-0.073	0.423	0.158	0.245	0.526	2.741	1.061	2.955	0.220	0.301	0.704	1.025	0.607	0.251
SFP	-0.018	-0.022	-0.011	-0.025	-0.010	-0.083	0.169	-0.014	-0.193	-0.001	-0.003	-0.043	-0.009	0.112
PB	-0.001	0.002	0.006	0.007	0.001	-0.004	-0.002	-0.005	0.000	-0.045	-0.022	-0.007	0.006	-0.070
SB	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.001	0.000	0.070
PW	0.001	0.008	-0.003	0.007	0.009	0.025	-0.004	0.022	0.014	0.010	0.020	0.064	0.005	0.169
TGW	0.016	0.018	0.001	0.004	0.017	0.025	0.009	0.026	0.006	-0.016	0.005	0.011	0.128	0.199

R SQUARE = 0.1607 RESIDUAL EFFECT = 0.9161, DFF - Days to 50 % flowering, PH - Plant height (cm), TN – Tiller number/plant, PN – Panicle number/plant,GN-Fertile grain number/panicle, CN –Chaff number/panicle, SN– Spikelet numbers/panicle, SFP – Spikelet fertility percentage, PL – Panicle length(cm), PB – Primary branching, SB-Secondary branching, PW- Panicle weight(gm), TGW- 1000 grain weight, YLD- Per se yield(gm)

percentage (0.4), primary branching (0.4), secondary branching (0.2), panicle weight (0.4) and thousand grain weight (0.14) while days to 50% flowering (-0.003) had negative indirect effect on grain yield and spikelet number had no indirect positive effect through any traits on grain yield (Table 4). Kumar and Saravanan (2012) and Minnie et al., (2013) reported similar results for days to maturity, number of productive tillers per plant, panicle length, fertile spikelet per panicle and spikelet fertility.

Cluster and principal component analysis

RILs developed from Reeta and WAB50-60 were statistically analyzed for diversity/similarity in their quantitative traits using cluster analysis by Ward method. The RILs were partitioned into three major clusters based on similarities in traits (Fig. 2). The major cluster I divided in to 2 sub clusters consisting of 7 RILs, major cluster 2 contains 10 sub clusters consisting of 167 RILs including parents and the major cluster 3 contains 3 sub clusters having 18 RILs. RILs falling in a particular cluster indicate their close relationship among themselves as compared to the other clusters.

Table 5. Eigen vectors and eigen values of the first four principal components.

Variables	PC 1	PC 2	PC 3	PC 4
DFF	0.05197	-0.06979	-0.11672	-0.04659
PH	0.21803	-0.00401	0.01155	-0.52482
TN	0.15632	-0.24414	0.35946	0.34792
PN	0.22282	-0.23996	0.32533	0.37901
PL	0.25758	0.02807	0.07344	-0.46288
GN	0.49748	0.09003	-0.02420	0.13970
CHN	-0.07100	0.58439	0.38478	0.03413
SN	0.43616	0.30623	0.12400	0.14305
SFP	0.26389	-0.48610	-0.36853	0.02564
PB	0.03270	0.29816	-0.49059	0.22829
SB	0.20672	0.30316	-0.33848	0.10240
PW	0.33312	0.10427	-0.18889	0.08726
TGW	0.18155	-0.01102	0.14308	-0.34805
YLD	0.33046	-0.07514	0.18592	-0.10138
Eigen Value	3.03945	2.09261	1.66382	1.34228
Variance (%)	21.71	14.947	11.884	9.5877
Cumulative variance (%)	21.71	36.657	48.541	58.1287

DFF - Days to 50 % flowering, PH - Plant height (cm), TN - Tiller number/plant, PN - Panicle number/plant, GN-Fertile grain number/panicle, CN-Chaff number/panicle, SN- Spikelet numbers/panicle, SFP-Spikelet fertility percentage, PL-Panicle length (cm), PB - Primary branching, SB-Secondary branching, PW- Panicle weight (gm), TGW- 1000 grain weight (gm), YLD- Per se yield (gm)

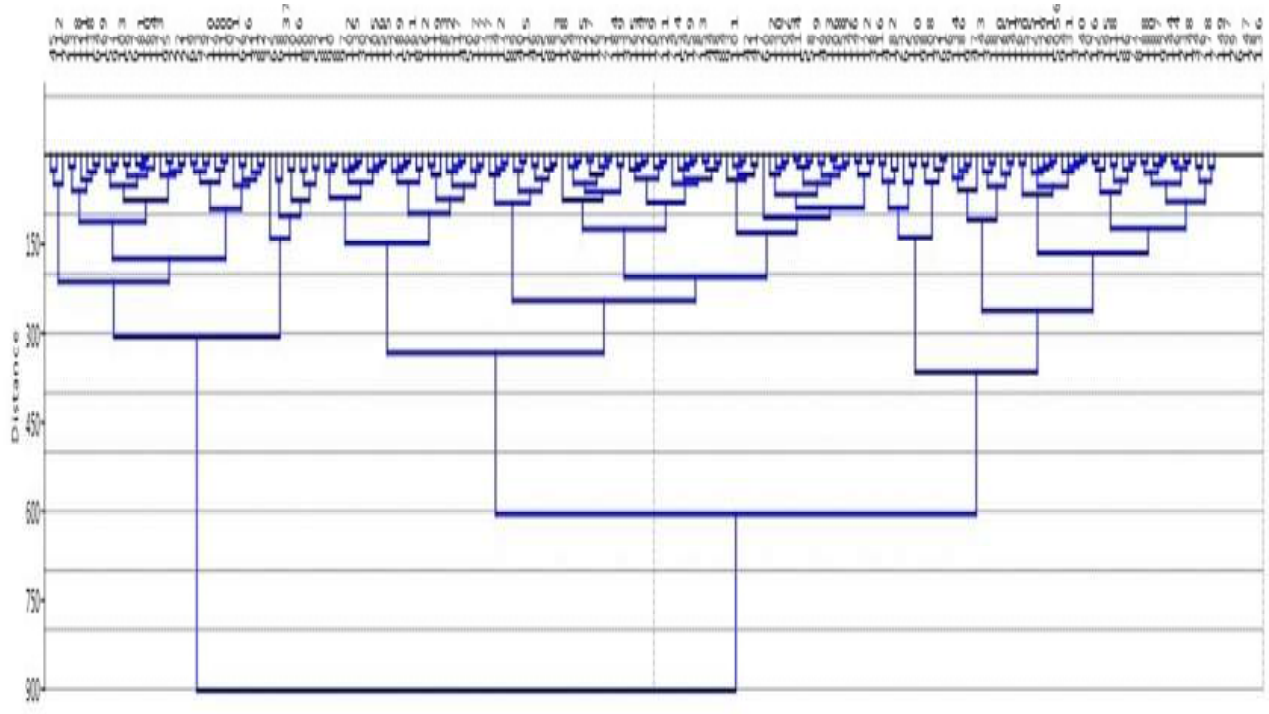

Fig. 2. Dendrogram showing genetic relationship between recombinant inbred lines and parents, Reeta and WAB56-50 based on Ward method. Numbers represent RILs.

Therefore, it could be expected that lines within a cluster were less genetically different with each other, and were diverse from the cultivars belonging to other clusters. These finding are in conformity with the earlier published results of (Singh et al., 2013; Pradhan et al., 2016a; Pradhan et al ., 2016b; Pandit et al ., 2016; Pandit et al., 2017). The genetic distance between the parents largely governs the variability spectrum generated in the segregating generation. Therefore, diverse RILs could be used in breeding programme for improvement of qualitative and quantitative traits. Principal component analysis was further used to establish the

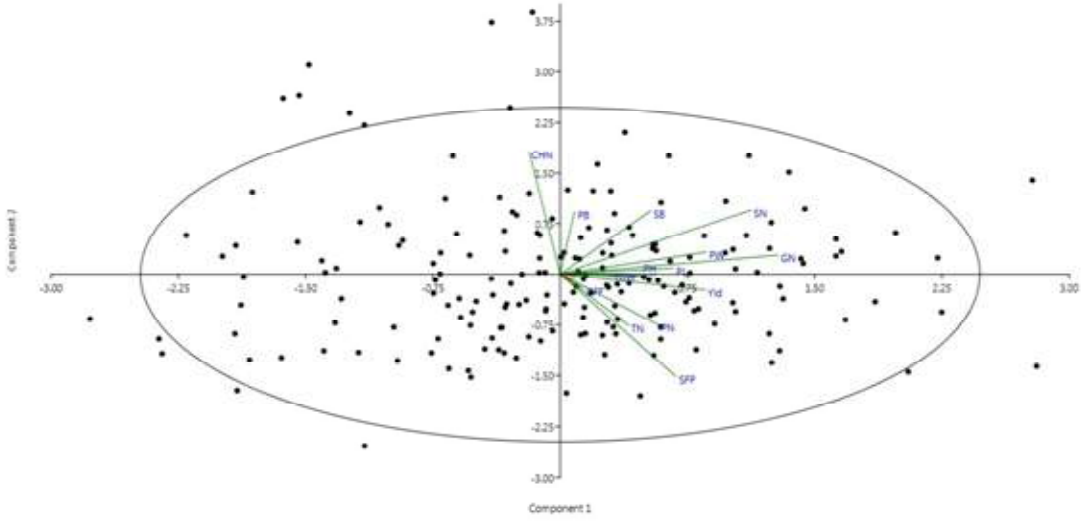

Fig. 3. Biplot graph showing unit variance scaling with imputation is used to calculate principal components. X and Y axis show principal component 1 and principal component 2 that explain 21.71% and 14.94%, respectively of the total variance.

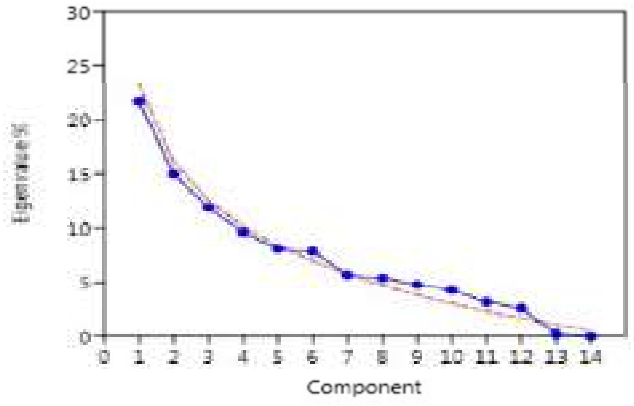

Fig. 4. Scree plot graph showing all the 14 principal components on X axis and percentage of Eigen value on Y axis.

patterns and interrelationships existing between the RILs and their quantitative traits (Fig. 3 and 4). The first four principal components explained a total of 58.12% variability in the all qualitative traits. The analysis of eigenvectors gave the information of qualitative traits for percentage of variation to the first four principal components, which were 21.71%, 14.94%, 11.88% and 9.58%, respectively (Table 5). Similarly, Singh et al., (2013) reported the first three principal components accounting for 62.72% of total variation among thirty five wild rice germplasm.

CONCLUSION

The accurate assessment of diversity at molecular level is important for crop improvement. Further, this will maximize the probability of transgressive segregation and accumulation of positive alleles from different cultivars. Microsatellites are considered to be appropriate for assessment diversity at molecular level, identification of cultivars and mapping of QTLs/ genes for different traits because of their ability to detect large numbers of discrete alleles accurately and efficiently. Present study indicated that high level of polymorphism exist between selected six genotypes. The polymorphic microsatellites in the study will be useful for identification of QTLs/ genes associated with yield traits such as grain number, 1000 grain weight, yield per se, etc using RIL mapping population developed from contrasting genotypes (PDKV Shriram and Heera, AC38562 and Pimpuidbasa, Reeta and WAB50-56). Further Principal component analysis indicated 58.12% of the total variation explained by first four principal components. PCA and cluster analysis complemented each other with some slight inconsistencies in terms of cluster composition. The separation and selection of RILs based on high heritability along with high genetic advance of traits make it easy for breeders to exploit their knowledge and skill in transgressive segregation breeding programme. High heritability along with high genetic advance as per cent of mean was recorded for chaff number, primary branching, panicle weight, thousand grain weight and total number of grains per panicle, while traits thousand grain weight, plant height and fertile spikelet per panicle, panicle length had positive direct effect and significant association with grain yield per plant at phenotypic and genotypic levels. Therefore, selections of segregating lines with desirable traits are to be effective in accumulation of favourable genes for bringing together into the common genetic background of cultivated indica rice (*Oryza sativa* L.)

REFERENCES

Al-Jibouri HA, Miller PA and Robinson HF (1958). Genotype and environmental variances in an upland cotton cross of inter specific origin. Agron. J. 50: 633-637

Anandan A, Mahender A, Kumari Rupa, Pradhan SK, Subudhi HN, Thirugnanakumar S and Singh ON (2016). Appraisal of genetic diversity and population structure in assorted rice genotypes for early seedling vigour trait linked markers. Oryza 53(2): 113-125

Behera L, Patra BC, Sahu RK, Nanda A, Sahu SC, Patnaik A, et al. (2012). Assessment of genetic diversity in medicinal rice using microsatellite markers. Aust. J. Crop Sci. 6(9): 1369-1376

Besufikad E, Sentayehu A, Abush T and John B (2016). Character correlation and path analysis for yield and yield related components in soybean (Glycine max L.) genotypes. Int. J. Sci. Appl. Res. 3(12): 97-104

Bhadru D, Tirumala R, Chandra M and Bharathi D (2012). Genetic variability and diversity studies in yield and its component traits in rice (*Oryza sativa* L.). SABRAO J. Breed. Genet. 44: 129-137

Dewey DR and Lu KH (1959). A Correlation and path coefficient analysis of components of crested wheat grass production. Agron. J. 52: 515-518

Fisher RA (1936). The use of multiple measurements in taxonomic problems. Annal. Eugen. 7: 179-188

Gangashetty P, Salimath P and Hanamaratti N (2013). Genetic variability studies in genetically diverse non-

basmati local aromatic genotypes of rice (*Oryza sativa* L). Rice Genom. Genet. 4 (2): 4-8

Iftekharuddaula KM, Badshah MA, Hassan MS, Bashar MK and Akter K (2001). Genetic variability, character association and path analysis of yield components in irrigated rice (*Oryza sativa* L.). Bangladesh J. Pl. Breed. Genet. 14(2): 43-49

Kayvan A, Mohammad H and Ejatollah F (2007). Correlation and path coefficient analysis for some yield-related traits in rice genotypes (*Oryza sativa* L.). Asian J. Plant Sci. 6 (3): 513-517

Khedikar VP, Bharose AA, Sharma D and Khillare AS (2004). Path coefficient analysis of yield components of scented rice. J. Soils and Crops. 14: 198-210

Kumar PS and Saravanan K (2012). Genetic variability, correlation and path analysis in rice (*Oryza sativa* L.). Int. J. Curr. Res. 4(9): 082-085

Kumar A, Reddy BVS, Ramaiah B, Sahrawat KL and Pfeiffer WH (2013). Gene effects and heterosis for grain iron and zinc concentration in sorghum [*Sorghum bicolor* (L.) Moench]. Field Crops Res. 146: 86-95. 10.1016/j.fcr.2013.03.001

Meier P (1954). Analysis of simple lattice designs with unequal sets of replications. J. American Stat. Assoc. 49(268): 786-863

Minnie CM, Reddy TD and Raju CS (2013). Correlation and path analysis for yield and its components traits (*Oryza sativa* L.). J. Res. ANGRAU 41(1): 132-134

Mohanty SK, Nanda A, Sahu RK, Panda RS, Behera L, Jena M, Sahu SC, Mohapatra SL and Mohapatra T (2017). Identifcation of novel quantitative trait loci associated with brown planthopper resistance in the rice landrace Salkathi . Euphytica 213: 38. DOI 10.1007/s10681-017-1835-2

Murray HG and Thompson WF (1980). Rapid isolation of high molecular weight DNA. Nucleic Acids Res. 8: 4321-4325

Oad FC and CruzSta P (2002). Rice varietal screening for ratoonability. Pakistan J. Appl. Sci. 2: 114-9

Pradhan SK, Nayak DK, Pandit E, Behera L, Anandan A, Mukharjee AK, Lenka S and Barik DP (2016a). Incorporation of bacterial blight resistance genes into lowland rice cultivars through marker assisted backcross breeding. Phytopathology 106: 710-718

Pradhan SK, Barik SR, Sahoo A, Mohapatra S, Nayak DK, Mahender A, Meher J, Anandan A and Pandit E (2016b). Population structure, genetic diversity and molecular marker-trait association analysis for high temperature stress tolerance in rice. PLoS ONE11(8): e0160027

Pandit E, Tasleem S, Barik S.R., Mohanty D. P., Nayak D.K., Mohanty SP, Das S and Pradhan SK (2017). Genome-Wide Association Mapping Reveals Multiple QTLs Governing Tolerance Response for Seedling Stage Chilling Stress in indica Rice. Frontier in Plant Science 8:552. doi: 10.3389/ fpls.2017.00552

Samonte SOPB, Wilson LT and McClung AM (1998). Path analyses of yield and yield-related traits of fifteen diverse rice genotypes. Crop Sci. 38: 1130-1136

Sarawgi AK, Rastogi NK and Soni DK (1997). Correlation and path analysis in rice accessions from Madhya Pradesh. Field Crops Res. 52: 161-167

Singh A, Singh B, Panda K, Rai VP, Singh AK, Singh SP, Chouhan SK, Rai V,Singh PK and Singh NK (2013). Wild rice of Eastern Indo-Gangetic plains of India constitutes two sub populations harbouring rich genetic diversity. Plant Omics J. 6(2): 121-127

Smith LA, Wang X, Peixoto AA, Neumann EK, Hall LM and Hall JC (1996). A Drosophila calcium channel alpha1 subunit gene maps to a genetic locus associated with behavioral and visual defects. J. Neurosci. 16(24): 7868-7879

Su CC, Zhai HQ, Wang CM, Sun LH and Wan JM (2006). SSR mapping of brown planthopper resistance gene Bph9 in Kaharamana, an indica rice (*Oryza sativa* L.). Acta. Genet. Sin. 33:262-268

Sun L, Su C, Wang C, Zhai H and Wan J (2005). Mapping of a major resistance gene to the brown planthopper in the rice cultivar Rathu Heenati. Breed. Sci. 55: 391-396

Sun LH, Wang CM, Su CC, Liu YQ, Zhai HQ and Wan J (2006). Mapping and marker-assisted selection of a brown planthopper resistance gene bph2 in rice (*Oryza sativa* L.). Acta. Genet. Sin. 33:171-723.

Surek H and Beser N (2003). Correlation and path coefficient analysis for some yield-related traits in rice (*Oryza sativa* L.) under stress conditions. Turk. J. Agric. For. 27: 77-83

Virk PS, Zhu J, Newbury HJ, Bryan GJ, Jackson MT and Ford-Lloyd BV (2000). Effectiveness of different classes of molecular marker for classifying and revealing variation in rice (*Oryza sativa L.*) germplasm. Euphytica 112: 275-284

Ward JH (1963). Hierarchical Grouping to Optimize an Objective Functional Journal of the American Statistical Association 58: 236-244

Williams WA, Jones MB and Demment MW (1990). A concise table for path analysis statistics. Agron. J. 82: 1022-1024

Wright S (1921). Correlation and Causation. J. Agric. Sci. 20: 557-587

Yang H, Ren X, Weng Q, Zhu L and He G (2002). Molecular mapping and genetics of a rice brown planthopper (*Nilaparvata lugens* Stal.) resistance gene. Hereditas 136: 39-43