International Journal of Agriculture Sciences
ISSN: 0975-3710&E-ISSN: 0975-9107, Volume 8, Issue 53, 2016, pp.-2631-2637.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000217

Research Article

ASSESSMENT OF GENETIC DIVERSITY OF WEEDY RICE (*Oryza sativa f. spontanea*) IN RELATION TO RICE RELATIVES OF ODISHA, INDIA

NGANGKHAM UMAKANTA, KATARA JAWAHARLAL, PARIDA MADHUCHHANDRA, MOHANTY MONALISA, PARAMESWARAN C., KUMAR AWADHESH, PATRA BHASKAR CHANDRA, SINGH ONKAR NATH AND SAMANTARAY SANGHAMITRA*

National Rice Research Institute (Formerly CRRI), Cuttack, 753006, Odisha, India
*Corresponding Author: Email-smitraray@gmail.com

Received: August 19, 2016; Revised: September 17, 2016; Accepted: September 18, 2016; Published: November 01, 2016

Abstract- Weedy rice is one of the most important noxious weed that infests commonly in the direct seeded rice fields worldwide. As weedy rice is conspecific form of cultivated rice, it is very difficult to discriminate it from the cultivars at its early growth stages, which are also difficult to control. Moreover, as the emergence of weedy rice is still obscure, the present study aimed at assessing the genetic diversity of seventy-five weedy rice collected from different regions of Odisha, India along with fifteen wild rice, three landraces and three popular cultivars grown in same field or in adjacent regions to understand the origin of weedy rice using SSR markers. A total of 14 SSR markers produced on an average of 4.12 alleles across 14 loci amplified. The polymorphic information content (PIC) ranged from 0.292 (RM 282) to 0.54 (RM 339) in all loci with an average of 0.48. The percentages of polymorphic loci were found to be 100%, 85.71% and 57.14% in weedy rice, wild rice and cultivars respectively, with the average of 80.95%. The dendrogram generated by UPGMA-based analysis clustered all the 96 individuals in three major groups with a similarity coefficient of 0.67. Principal coordinate analysis (PCoA) showed 30.27% of all genetic variation was clustered in first two components in the analysis. Similar to cluster analysis, structure analysis also clustered all the rice genotypes into three major groups comprising of 96 genotypes (weedy rice, weedy rice, wild rice, landraces and cultivars and admixtures respectively). This is the first preliminary report of weedy rice diversity in India, which shows the evolution of weedy rice as a complex phenomenon. This needs further studies for better understanding of weedy rice evolution in India using more number of markers along with genes responsible for domestication of rice.

Keywords- Weedy rice, Red rice, Genetic diversity, Simple Sequence Repeat (SSR), genetic relationship

Citation: Ngangkham Umakanta., et al., (2016) Assessment of Genetic Diversity of Weedy Rice (*Oryza sativa f. spontanea*) in Relation to Rice Relatives of Odisha, India . International Journal of Agriculture Sciences, ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 8, Issue 53, pp.-2631-2637.

Academic Editor / Reviewer:

Introduction

Rice (*Oryza sativa* L.) is one of the most important cereal food crops for more than one-half of the world population and provides 50–80% of daily calorie intake [1]. Considering the world population growth rate and food security with diminishing arable land in the world, an estimate of 40 % increase in rice production will be needed by 2030 [1]. One of the major constraints/bottlenecks to attain the potential yield target of rice cultivars is the contamination of cultivated rice with weedy rice. Infestation of rice production areas with weedy rice causes yield reduction up to 80% [2, 3] and affects the quality of rice grains [4]. Weedy rice is distributed broadly across rice-growing regions, especially where direct seeding and intensive production prevails, and weedy rice seems to possess a wide variation in its biological characteristics [5, 6].

Weedy rice is a conspecific form of cultivated rice (*Oryza sativa* L.) that infests rice production areas worldwide [2]. Weedy rice typically grows naturally in and around rice fields and it is an aggressive competitor that spreads rapidly and drives down the quality of the rice harvest and specifically its high propensity to shattering ensures its continuous propagation in the field [7-9]. Some accessions of weedy rice are similar to wild rice, which also have a red pericarp, black hull, long awn, light seed weight, strong seed dormancy and easy seed shattering [6]. Moreover, the morphological and physiological similarities exist between weedy rice and cultivated rice, it is difficult to detect in its early growth stages which is difficult to control with herbicides [10]. There is very high level of genetic variability and plasticity observed within and among weed populations [11] as the growth of weedy rice varies considerably among different biotypes due to differences in plant height, tillering, or leaf-producing capacity. Besides, some of the important traits like early seed shattering, variable seed dormancy and high seed persistence in the soil varies in some of the weedy rice. This enables weedy rice to infest a wide range of diverse habitats and it competes aggressively with cultivated rice. Thus, a full understanding of the genetic diversity of weeds is important in effective management of weedy rice [12] which would give an idea in elucidating the origin and evolutionary processes of weedy rice; this would be very much helpful for designing effective management strategies to control the weedy rice [13].

Since differences in opinions on origin and evolution of weedy rice were expressed in different countries like US, China and other developing countries, genetic analyses suggest a very close relationship of weedy rice to domesticated rice, suggesting an origin by dedomestication from feral crop strains [2, 14, 15 & 16], whereby selection favours reversions of domestication traits to form characterizing wild species [17]. Besides, based on morphological and genetic analysis of weedy rice, it has been elucidated that evolution of weedy rice was due to hybridization between the wild rice (*Oryza rufipogon*) and the cultivated rice [18, 19 & 20], or hybridization between the *indica* and *japonica* [21, 22] or directly evolved from hybrid rice [23]. With reports to date being meager on the evolution of weedy rice in India, the present study focuses on the evaluation of genetic

diversity in weedy rice collected from Odisha along with improved cultivars, landraces and wild rice using SSR markers to understand the origin and evolution of weedy rice.

MaterialsandMethods

Plant Material

The plant samples used in this study includes seventy five weedy rice collected from different locations of t Odisha, India; fifteen wild rice (six accessions of *Oryza rufipogon*, nine accessions of *Oryza nivara*), six cultivated rice including three each of landraces and high yielding cultivars. Seed samples of weedy rice were collected from rice fields of different locations of Odisha state, India during 2012. The samples have been collected with a distance of at least 1 km from each other and each sample was regarded as members of one population. The samples were grown in the research field of NRRI, Cuttack for consecutive three seasons to observe the stability and segregation. The samples found to be stable and uniform, were selected for the present studies. The seeds were sown in the nursery and 21 days old seedlings were transplanted in the research field of the NRRI, Cuttack following the standard agronomic practices.

Genomic DNA isolation

The young leaf from each samples were collected and frozen in liquid nitrogen and stored at -80°C. The genomic DNA was isolated as per the standardized CTAB method (cetyltrimethyl ammonium bromide) [24]. The isolated genomic DNA was quantified and further diluted to 30 ng/µl for use in PCR amplification.

SSR markers and PCR amplification

A set of 14 SSR markers were taken for genotyping of 96 individuals. The details information of the primer pairs for 14 SSR markers used in the present study is given in [Table-1]. The PCR reaction was carried out in a 10µl reaction mixture containing 30ng of genomic DNA, 10 pmol of both forward and reverse primers, 0.2mM of dNTP mix, 1U of *Taq* polymerase (Bangalore genie, India), 1X Taq buffer with 1.5 mM of $MgCl_2$. The PCR amplification was performed in a Thermal Cycler (Eppendorf, USA) using the following cycling program; Initial denaturation at 94°C for 5 min followed by 35 cycle of denaturation at 94°C for 45 sec, annealing at 55°C for 30 sec and polymerization at 72°C for 1 min and completed the reaction by final extension at 72°C for 10 min. The amplified PCR products were resolved in a 3.5% Metaphor agarose gel stained with ethidium bromide. Gel pictures were visualized using Gel documentation system (Alpha Imager®, USA).

Table-1 *Details of SSR markers used in this study*

Sl no.	Primer ID	Forward Sequences(5' to 3')	Reverse Sequences(5' to 3')	SSR motif
1	RM431	GCTTGCTTGTATCTGCATTGGTAGG	GGGATGATCCACTCTCTGTTTGG	(AG)16
2	RM495	ATGATGATGGACGACGACAACG	TGAATCCAAGGTGCAGAGATGG	(AGC)7
3	RM580	GAAATGGACTCGCTCCTAAACTGG	ACGAACTAGAGCATGGGCACTCC	(AAG)19
4	RM154	GACGGTGACGCACTTTATGAACC	CGATCTGCGAGAAACCCTCTCC	(AG)21
5	RM475	CCTCACGATTTTCCTCCAAC	ACGGTGGGATTAGACTGTGC	(TATC)8
6	RM71	CTAGAGGCGAAAACGAGATG	GGGTGGGCGAGGTAATAATG	(ATT)10T(ATT)4
7	RM166	GGTCCTGGGTCAATAATTGGGTTACC	TTGCTGCATGATCCTAAACCGG	(T)12
8	RM514	CTTCTCAGATTGATCTCCCATTCC	GGGAGAGAGGAAGAAGACAAGG	(AC)12
9	RM545	CCTTCCCTGAAAGTATTCGTTCTCC	GAGAACGTCTTCATTGGATGTTCC	(AG)30
10	RM471	AGAAATGGATCGGACTGAACATGC	AGACACTCGGACGCACAAGC	(AG)12
11	RM551	CTTACTCCATTGGGCTGGAACC	TGTAGGGTGGTAAGAGATCCACTCC	(AG)18
12	RM153	CCTCGAGCATCATCATCAGTAGG	TCCTCTTCTTGCTTGCTTCTTCC	(AAG)9
13	RM480	TGGTACTCACCATGCAAGTAGAACG	ATGCTCAAGCATTCTGCAGTTGG	(AC)30
14	RM249	GGCGTAAAGGTTTTGCATGT	ATGATGCCATGAAGGTCAGC	(AG)5A2(AG)14

Scoring of SSR amplified loci and Statistical analysis

The amplified SSR DNA bands denoting different alleles were scored as binary data whether present (1) or absent (0). The binary matrix of the SSR marker scores were used to assess the genetic relationships analysis by estimating genetic distance and similarity coefficients. Then, the similarity matrices were subjected to cluster analysis as per the unweighted pair-group method with arithmetic averages (UPGMA), and developed dendrogram using NTSYS-PC version 2.0 [25]. Genetic diversity parameters such as number of alleles per locus, allele frequency, heterozygosity, gene diversity and polymorphic information index (PIC) was estimated using the program POWERMARKER Ver3.25 [26]. The Shannon's diversity index (I), Fixation index and private alleles were calculated to estimate the level of genetic diversity using GenAlEx 6.502 [27]. The program STRUCTURE version 2.3.4 [28] was implemented to determine the population structure and to identify the number of clusters (K). The number of clusters was run from K = 1 to K = 10, with 5 independent replications per K using the admixture model and correlated allele frequencies, a 200,000 burn-in period and 200,000 MCMC. Then, ΔK was used to determine the optimal K according to the method described by Evanno et al. [29] using STRUCTURE HARVESTER programme [30]. Approximations of posterior distributions are obtained using Markov chain Monte Carlo (MCMC) methods. The optimum value of K was then determined at the peak of ΔK. For each value of K, STRUCTURE generates a Q-matrix (QST) that lists the estimated membership coefficients for each genotype in each subgroup. An individual having more than 80% of its genome fraction value under a particular K subgroup was assigned to that subgroup. The PCoA (Principle coordinate analysis) was computed using the binary data of SSR markers through GenAlEx 6.502 [27]. The values of the eigenvectors obtained were plotted in a scatter graph taking the first principal component and the second principal component as the axes. A hierarchical analysis of molecular variance (AMOVA) of weedy rice populations was calculated to partition genetic diversity using GenAlEx 6.502 [27].

Results and Discussion

Weedy rice, also called as red rice because of its red pericarp, has been reported in almost all rice-growing regions in the world [31]. Weedy rice is a notorious weed in the paddy field, which hinders almost 100% grain yield. Chemical control measures in conventional rice cultivars are not an easy option to manage weedy rice, simply because of the similar physiological and phenotypic traits between weedy rice and cultivated rice. Since, evolution of the weedy rice is not completely understood, a preliminary study was undertaken to evaluate the genetic diversity of weedy rice lines found in the state of Odisha, India using SSR markers from different chromosomes of rice.

Genetic diversity of the weedy rice

Genotyping analysis using selected 14 SSR markers produced a total of 43 alleles and the number of alleles per marker ranged from 2 to 6 with an average of 3.07 alleles per locus [Table-2]. All the 14 SSR markers used here were found to be

polymorphic. This finding corroborated the earlier reports in assessing the genetic diversity of 208 weedy rice of China using SSR markers [31]. The percentages of polymorphic loci were found to be 100%, 85.71% and 57.14% in weedy rice, wild and cultivars respectively with the average of 80.95%. Thus, the selected 14 SSR markers were found to be robust which were used for diversity analysis. Among the 14 SSR markers used, the highest number of alleles (6) was detected with RM580 followed by 5 with RM153 and 3 with RM480, RM249, RM495, RM471, RM545, RM551, and RM71; the lowest of 2 alleles were detected in four SSR markers such as RM431, RM154, RM475 and RM66. In the present study, the maximum (0.73) and minimum (0.15) PIC value was recorded in RM580 and RM66 respectively with an average PIC value of 0.407. This result is similar to Courtois et al. [32] who reported the PIC value in the range of 0.16 to 0.78 with an average of 0.49 in genetic diversity study of European rice germplasm. Similarly, the mean PIC value of 0.48 and 0.4214 were observed in 416 rice accessions of China and weedy rice of Guandgong provinces in China respectively [31, 33]. This indicated the similar pattern of evolution and fixation of alleles of weedy rice in Indian vis-a-vis Chinese weedy rice collections. Besides, the PIC value revealed from 96 genotypes used in the present study was more diverse due to differences in origin and speciation.

The observed heterozygosity (Ho) in our analysis ranged from 0 to 0.16 with an average of 0.05 which is found low. The fixation index also varied between 0.656 and 1.00 with an average of 0.909 [Table-2]. This low heterozygosity might be due to autogamous nature of rice crop [34] and other possible reasons are either due to less segregation or due to selection of stable progenies in our experimental plot. The private alleles were defined as the allele found only in a population, which may be considered as unique allele to that population [35, 36]. In our study, a private allele was detected in RM166 only in weedy rice with the frequency of 12.8%, which is much higher as compared to the Jiang et al. [37]. Obtaining such high private allele in weedy rice showed the differentiation of weedy rice from other rice population used in this study, which confirms the self-fertilization mating system in weedy rice collections of Odisha, India corroborated the finding of Song et al. [38]. Contrastingly, moderate level of Ho value (>0.13) indicate an increased degree of heterozygosity resulted from outcrossing in weedy rice populations [38]. Expected heterozygosity or Gene diversity (He) computed in this analysis varied from 0.16 (RM66) to 0.77 (RM287) with the average of 0.47 which was similar to US and Chinese rice accession panel [33, 39] and also in several other recent past studies [37, 38, 40, 41, 42, & 43]. Interestingly, the average values of genetic diversity in our study shows much higher than the weedy rice populations sampled from China [38]. Thus, the considerable higher level of genetic diversity in weedy rice lines of Odisha indicates a complicated type of origin of weedy rice in this region of India. The probable reason of such higher diversity might be due to either adoption of direct seeded rice growing technique in this region or reduced weed control practices owing to limited human labor input. Moreover, recent changes in farming practices and cultivation methods along with less weed management directly or indirectly involves in re-emergence and divergence of weedy rice.

Table-2 *Microsatellite diversity at the fourteen SSR loci in 96 genotypes*

Markers	Major allele frequency	Genotype No.	Sample size	Allele no.	Gene diversity	Heterozygosity	PIC	F	I
RM153	0.5000	7	96	5	0.5347	0.1052	0.4267	0.884	0.689
RM480	0.6808	4	96	3	0.4379	0.1276	0.3465	0.755	0.626
RM431	0.5238	2	96	2	0.4988	0	0.3744	1.000	0.692
RM154	0.8085	3	96	2	0.3096	0.1063	0.2617	0.656	0.488
RM249	0.4555	5	96	3	0.6242	0.0888	0.5462	0.867	0.686
RM495	0.6979	5	96	3	0.4516	0.0833	0.3914	0.884	0.562
RM580	0.3206	11	96	6	0.7750	0.1630	0.7399	0.864	0.684
RM514	0.6947	5	96	3	0.4613	0.0842	0.4066	0.818	0.562
RM475	0.8354	2	96	2	0.2749	0	0.2371	1.000	0.447
RM471	0.6600	4	96	3	0.5055	0.0133	0.4525	1.000	0.472
RM545	0.5116	3	96	3	0.5970	0	0.5184	1.000	0.675
RM551	0.5106	3	96	3	0.5461	0	0.4444	1.000	0.690
RM71	0.7157	3	96	3	0.4445	0	0.3996	1.000	0.500
RM166	0.9076	2	96	2	0.1675	0	0.1535	1.000	0.305
Mean	0.6302	4.2142	96	3.0714	0.4735	0.0551	0.4071	0.909	0.577

PIC: Polymorphism information content; F: Fixation Index; I: Shannon's Information Index

Genetic relationship of weedy rice

In order to assess the genetic relationships among the weedy rice, wild rice, land races and cultivated rice of Odisha, the UPGMA method was deduced to generate a dendrogram which showed a cluster of 96 individuals into three main groups named as I, II, and III with a similarity coefficient of 0.67 [Fig-1]. The main group I comprising 67 genotypes further divided into two subgroups such as A and B. The main group II included 25 genotypes formed two subgroups; C and D. A similar analysis from 206 weedy rice collections had only two major groups in the dendrogram [31]. This indicates that mixtures of weedy rice along with cultivars are present in Odisha, India rather forming distinct groups. The genetic similarity coefficient of the whole 96 samples varied from 0.60 to 1. The largest major group B was further divided as B-I, B-II and B-III clusters. The B-I cluster comprised of 32 genotypes as the largest cluster, followed by B-II with 25 genotypes and B-III with 3 genotypes. The major group A with 7 genotypes is further divided as A-I (3) and A-II (4). Interestingly, the cluster A-I comprised of three genotypes including IR64, Swarna and a weedy rice line, PM113. As expected, *indica* cultivars IR64 and Swarna have greater genetic similarity coefficient of 0.97 as compared to the rest of the samples studied. However, presence of a weedy rice line, PM113 indicates that closer evolution with these two cultivars or a non causal random association. This can be resolved by utilizing more number of SSR markers for dendrogram analysis. Similarly, mixtures of cultivars along with weedy rice were also reported in Chinese weedy rice collections [31]. Another improved cultivar, Pooja was found along with weedy rice and O. *nivara* lines in the Cluster B-II which indicated that cultivar Pooja has diverse genetic components as compared to other improved cultivars like IR64 and Swarna but more similar to O. *nivara*, wild rice. The largest cluster B-I consisting of four diverse genotypes *viz*. 26 weedy rice (34.6%), 2 landraces, 2 O. *nivara* and 2 O. *rufipogon*. This indicates the diversification of weedy rice lines into distinct groups after their evolution. The smallest genetic distance was observed between the two weedy rice, PM76 and PM78 within the cluster B-I. In the main group II, the subgroup C is different from the rest due to the presence of only wild rice, O. *rufipogon*. The wild rice in this group formed a relatively independent clade from the rest of the rice individuals. Separate clustering of O. *rufipogan* might be due to the perennial nature of this species which suggests the role of O. *rufipogan* contributing to evolution of weedy

rice in India might be less as compared to other wild rice, O. *nivara*. Interestingly, main group III contains only four weedy rice exhibiting largest genetic distance with respect to other populations with a genetic similarity coefficient of 0.6. Based on this fourteen SSR markers data analysis, 75 weedy rice used in the present study were distributed in all the clusters except subgroup C. This result suggests a complex and unclear evolutionary process of weedy rice in Odisha, India.

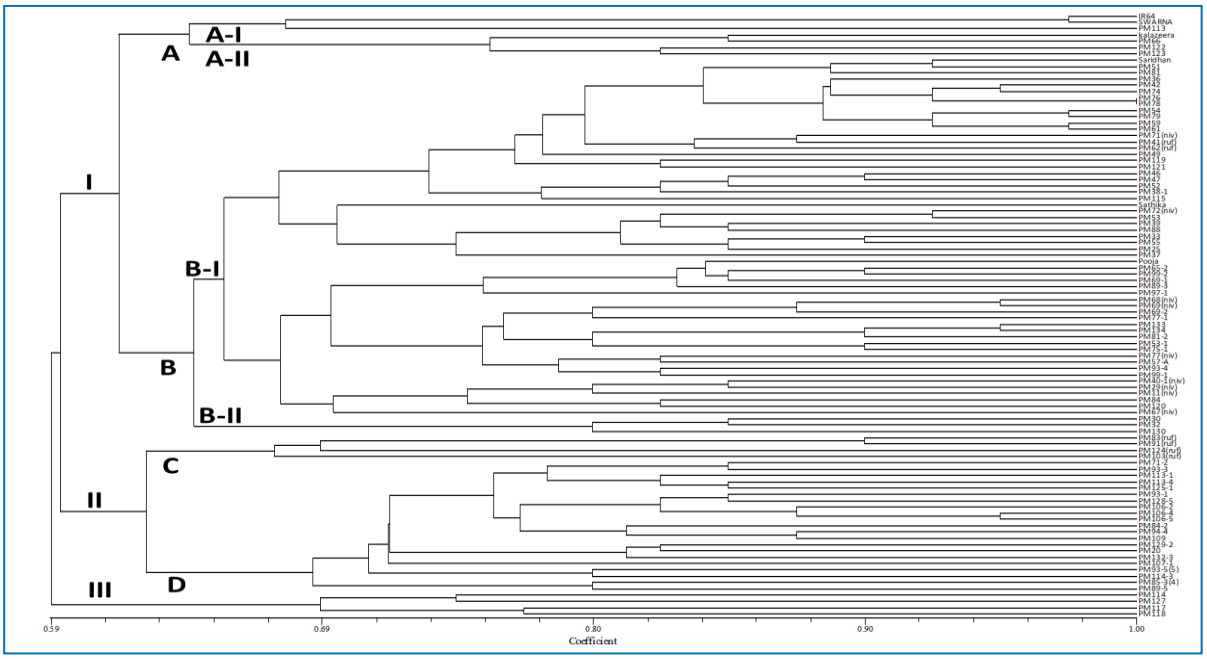

Fig-1 Dendrogram of weedy, cultivar and wild rice using UPGMA based on Nei's similarity coefficients using NTSYS-pc 2.21 software. Niv and ruf denoted O. nivara and O. rufipogan respectively.

To estimate genetic relationships of the weedy rice with the other rice individuals, PCoA was inferred based on the 14 SSR markers data matrix to characterize the subgroups of the 96 rice individuals. A scatter plot generated from the PCoA analysis had shown that the first two components accounted for about 30.27% of all genetic variation [Fig-2] These scatter plots showed a complex genetic relationship among the weedy rice, wild and cultivated rice. Such pattern of grouping of weedy rice in PCoA analysis was also reported by Zhang et al. [31]. Most of the weedy rice individuals were dispersed across the positive and negative loads of the first principal component suggesting relatively complex sources and origins of weedy rice in Odisha collections and mechanism of evolution is distinct for different weedy rice lines. This result is concurrent with the result of the UPGMA-based analysis. Thus, judging from the genetic diversity patterns of this study, the origin of some of the weedy rice of Odisha is probably through the hybridization between the wild rice and rice cultivars cultivated in the nearby areas by the farmers.

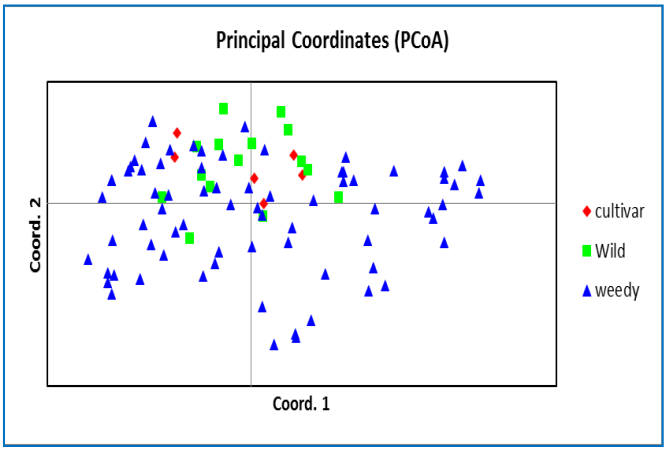

Fig-2 Two-dimensional PCoA display of 96 rice individuals based on 14 SSR loci. Coord 1 and Coord 2 represent first and second coordinates, respectively. The two PCA axes accounted for 17.14 and 13.13 % of the genetic variation among populations.

Genetic structure of weedy rice

The genetic architecture of diverse germplasm lines can be precisely estimated by assessing the STRUCTURE of the population using molecular markers such as SSRs or SNPs etc. [44, 45 & 46]. Population structure of the 96 genotypes was analyzed using Bayesian based approach. The peak plateau of adhoc measure ΔK was found to be K=3 [Fig-3], which indicates the grouping of entire 96 genotypes into three subgroups (SGs). Population structure analysis in different rice diversity panel has already indicated the existence of two to eight subpopulations in rice [34]. Classification of rice genotypes to the subgroups based on ancestry threshold varies and ranges from the threshold of 60% [47] to 95% [48]. Our result showed a stringent fixed threshold value of 80% to identify genotypes belonging to a specific subgroup, while rest below the threshold were categorized as admixtures [Fig-4]. Specifically, SG-1 consisted of 9 genotypes of which all of them were weedy rice. In SG-2, there were 32 genotypes mostly from wild rice, O. *nivara* and weedy rice. The SG-3 solely comprised of twenty four genotypes which includes all cultivars used in this study, 17 weedy rice and few wild rice (O. *nivara* and O. *rufipogon*). The remaining 31 genotypes showing below threshold value retained as admixture holding with maximum number of weedy rice (21) and few of wild rice; O. *nivara* (4) and O. *rufipogon* (4) rice and landraces (2). This analysis shows the complicated origin and evolution of weedy rice of Odisha collections. Likewise, the analysis of molecular variance (AMOVA) showed 86% variation among all the individuals whereas 9% variation was observed among population and 6% within individuals (P<0.001) [Fig-5]; [Table-4]. This result shows an extensive genetic diversity among weedy rice populations. This study also reveals a major genetic difference between individual rice varieties especially among the weedy rice whereas a small proportion of diversity occurred among the population of weedy, wild and cultivars of rice. Therefore, the main contribution to the genetic variation in the genotypes taken for the study was due to variation within genotypes while differences among populations had only 9% contribution to the total genetic variation. This small genetic difference between the populations could be possibly due to presence of high percentage of admixture in the genotypes selected. These results were similar to the earlier reports in Indian rice varieties [49, 50].

Table-3 *Population structure group of genotypes based on Inferred ancestry values*

Sample No.	Genotypes	Inferred Ancestry			Structure group	Subtype
		Q1	Q2	Q3		
1	IR64	0.017	0.012	0.971	SG3	cultivar
2	SWARNA	0.02	0.014	0.965	SG3	cultivar
3	PM 9	0.013	0.649	0.339	AD	cultivar
4	PM6	0.009	0.502	0.489	AD	cultivar
5	Pooja	0.067	0.011	0.922	SG3	cultivar
6	kalazeera	0.099	0.017	0.883	SG3	cultivar
7	PM 40-1	0.009	0.017	0.974	SG3	wild (nivara)
8	PM 67	0.155	0.781	0.064	AD	wild (nivara)
9	PM 68	0.013	0.955	0.032	SG2	wild (nivara)
10	PM 69	0.049	0.588	0.363	AD	wild (nivara)
11	PM 71	0.038	0.748	0.213	AD	wild (nivara)
12	PM 72	0.009	0.977	0.014	SG2	wild (nivara)
13	PM 77	0.014	0.977	0.008	SG2	wild (nivara)
14	PM 11	0.02	0.968	0.013	SG2	wild (nivara)
15	PM 29	0.031	0.401	0.568	AD	wild (nivara)
16	PM 41	0.008	0.284	0.709	AD	wild (rufipogon)
17	PM 62	0.01	0.265	0.725	AD	wild (rufipogon)
18	PM 83	0.013	0.069	0.917	SG3	wild (rufipogon)
19	PM 91	0.016	0.102	0.881	SG3	wild (rufipogon)
20	PM103	0.023	0.395	0.582	AD	wild (rufipogon)
21	PM 124	0.024	0.279	0.696	AD	wild (rufipogon)
22	PM 30	0.052	0.927	0.02	SG2	weedy
23	PM 32	0.277	0.648	0.075	AD	weedy
24	PM 33	0.014	0.878	0.108	SG2	weedy
25	PM36	0.009	0.977	0.014	SG2	weedy
26	PM 37	0.012	0.487	0.501	AD	weedy
27	PM 39	0.035	0.909	0.057	SG2	weedy
28	PM 42	0.007	0.977	0.016	SG2	weedy
29	PM 46	0.008	0.981	0.011	SG2	weedy
30	PM 47	0.005	0.978	0.017	SG2	weedy
31	PM 49	0.108	0.565	0.326	AD	weedy
32	PM 51	0.008	0.975	0.017	SG2	weedy
33	PM 52	0.009	0.977	0.014	SG2	weedy
34	PM 53	0.009	0.976	0.014	SG2	weedy
35	PM 54	0.009	0.954	0.037	SG2	weedy
36	PM 55	0.017	0.974	0.009	SG2	weedy
37	PM 59	0.009	0.962	0.029	SG2	weedy
38	PM 61	0.01	0.967	0.022	SG2	weedy
39	PM 66	0.01	0.042	0.948	SG3	weedy
40	PM 74	0.006	0.98	0.014	SG2	weedy
41	PM 76	0.006	0.959	0.034	SG2	weedy
42	PM 78	0.006	0.96	0.034	SG2	weedy
43	PM 79	0.128	0.815	0.056	SG2	weedy
44	PM 81	0.017	0.627	0.356	AD	weedy
45	PM 84	0.065	0.331	0.604	AD	weedy
46	PM 88	0.017	0.933	0.049	SG2	weedy
47	PM 25	0.146	0.834	0.02	SG2	weedy
48	PM 113	0.021	0.024	0.956	SG3	weedy
49	PM 114	0.775	0.016	0.209	AD	weedy
50	PM 115	0.016	0.975	0.009	SG2	weedy
51	PM 117	0.74	0.216	0.044	AD	weedy
52	PM118	0.053	0.148	0.799	AD	weedy
53	PM119	0.031	0.645	0.323	AD	weedy
54	PM120	0.059	0.761	0.18	AD	weedy
55	PM 121	0.007	0.981	0.011	SG2	weedy
56	PM 122	0.392	0.278	0.33	AD	weedy
57	PM 123	0.138	0.226	0.636	AD	weedy
58	PM 127	0.44	0.134	0.426	AD	weedy
59	PM 130	0.053	0.938	0.009	SG2	weedy
60	PM 133	0.067	0.923	0.009	SG2	weedy
61	PM 134	0.038	0.95	0.011	SG2	weedy
62	PM 38-1	0.009	0.978	0.013	SG2	weedy
63	PM 53-1	0.555	0.181	0.263	AD	weedy
64	PM 65-2	0.01	0.015	0.975	SG3	weedy
65	PM 69-1	0.008	0.055	0.937	SG3	weedy
66	PM 69-2	0.012	0.803	0.185	SG2	weedy
67	PM 71-2	0.025	0.007	0.967	SG3	weedy
68	PM 57-A	0.546	0.409	0.044	AD	weedy
69	PM 75-1	0.227	0.711	0.062	AD	weedy
70	PM 77-1	0.01	0.973	0.017	SG2	weedy

71	PM 81-2	0.278	0.688	0.034	AD	weedy
72	PM 84-2	0.014	0.04	0.947	SG3	weedy
73	PM 85-3(4)	0.926	0.022	0.052	SG1	weedy
74	PM 89-5	0.334	0.121	0.545	AD	weedy
75	PM 89-3	0.013	0.03	0.957	SG3	weedy
76	PM 93-1	0.018	0.044	0.938	SG3	weedy
77	PM 93-3	0.014	0.008	0.978	SG3	weedy
78	PM 93-4	0.194	0.749	0.057	AD	weedy
79	PM 93-5(5)	0.016	0.016	0.968	SG3	weedy
80	PM 94-4	0.016	0.012	0.972	SG3	weedy
81	PM 97-1	0.017	0.045	0.938	SG3	weedy
82	PM 99-1	0.262	0.34	0.397	AD	weedy
83	PM 99-2	0.015	0.024	0.96	SG3	weedy
84	PM 106-2	0.157	0.073	0.77	AD	weedy
85	PM 106-4	0.015	0.016	0.969	SG3	weedy
86	PM 106-5	0.017	0.015	0.968	SG3	weedy
87	PM 107-1	0.055	0.017	0.928	SG3	weedy
88	PM 109	0.012	0.018	0.97	SG3	weedy
89	PM 113-1	0.974	0.011	0.015	SG1	weedy
90	PM 113-4	0.975	0.009	0.016	SG1	weedy
91	PM 114-3	0.985	0.007	0.008	SG1	weedy
92	PM 125-1	0.957	0.01	0.032	SG1	weedy
93	PM 128-5	0.969	0.01	0.021	SG1	weedy
94	PM 129-2	0.981	0.009	0.01	SG1	weedy
95	PM 132-3	0.981	0.011	0.009	SG1	weedy
96	PM20	0.967	0.015	0.018	SG1	weedy

Based on the membership fractions, the accessions with the probability of >80% were assigned to corresponding subgroups with others categorized as admixture. SG: Subgroup, AD: Admixture

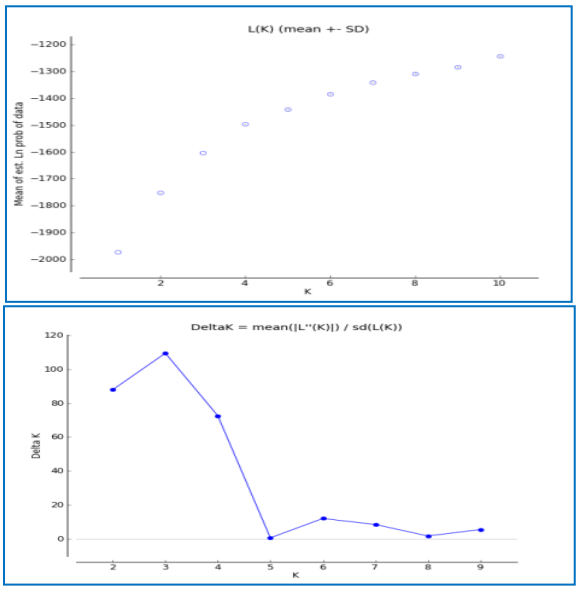

Fig-3 ΔK over five repeated runs of Structure simulations for 96 rice genotypes. The maximum of adhoc measure ΔK determined by structure harvester was found to be K = 3, which indicated that the entire population can be grouped into three subgroups.

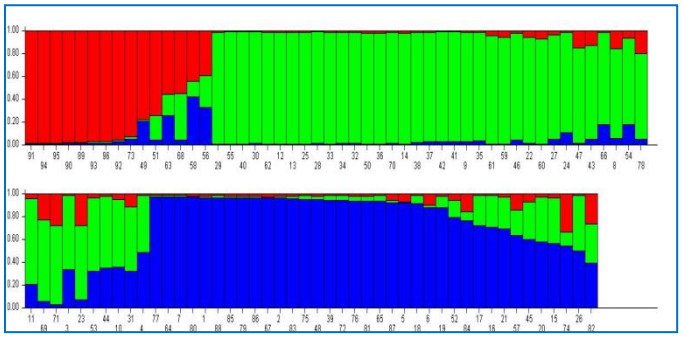

Fig-4 Population structure of 96 genotypes arranged based on inferred ancestry. Based on the membership fractions, the accessions with the probability of >80% were assigned to corresponding subgroups with others categorized as admixture. Each sample is represented by a single vertical line broken into K (K=3) colored segments, with lengths proportional to each of the K inferred clusters. Different color within group indicates the proportion of shared ancestry with other group, which has the same color with the admixture.

Sl no.	Primer ID	Forward Sequences (5' to 3')	Reverse Sequences (5' to 3')	SSR motif
1	RM431	GCTTGCTTGTATC TGCATTGGTAGG	GGGATGATCCACTCT CTGTTTGG	(AG)16
2	RM495	ATGATGATGGACG ACGACAACG	TGAATCCAAGGTGCA GAGATGG	(AGC)7
3	RM580	GAAATGGACTCGC TCCTAAACTGG	ACGAACTAGAGCATG GGCACTCC	(AAG)19
4	RM154	GACGGTGACGCA CTTTATGAACC	CGATCTGCGAGAAAC CCTCTCC	(AG)21
5	RM475	CCTCACGATTTTC CTCCAAC	ACGGTGGGATTAGAC TGTGC	(TATC)8
6	RM71	CTAGAGGCGAAAA CGAGATG	GGGTGGGCGAGGTA ATAATG	(ATT)10 T(ATT)4
7	RM166	GGTCCTGGGTCAA TAATTGGGTTACC	TTGCTGCATGATCCTA AACCGG	(T)12
8	RM514	CTTCTCAGATTGA TCTCCCATTCC	GGGAGAGAGGAAGAA GACAAGG	(AC)12
9	RM545	CCTTCCCTGAAAG TATTCGTTCTCC	GAGAACGTCTTCATT GGATGTTCC	(AG)30

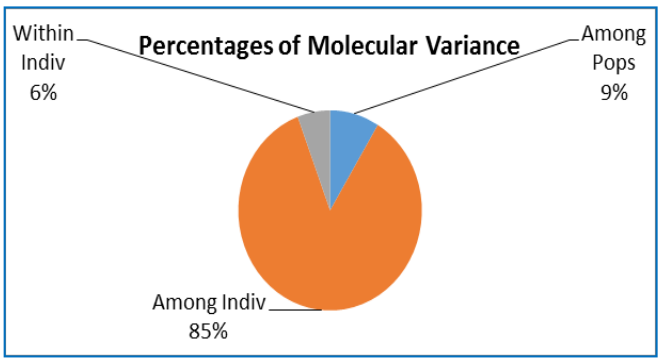

Fig-5 Partitioning of molecular variance among weedy rice, wild rice and rice cultivars using GenAlEx 6.502 software.

Table-4 *Analysis of molecular variance (AMOVA) of 96 rice individuals*

Source	df	SS#	MS^	Est.*	Var.%†	P Value‡
Among Pops	2	38.189	19.095	0.347	9	0.001
Among Individuals	93	655.373	7.047	3.409	86	
Within Individuals	96	22.000	0.229	0.229	6	
Total	191	715.563		3.985	100	

#SS, sum of squares.
^MS, mean square.
*Est., variance component.
†Var.%, percentage of variance.
‡Probability is based on 999 permutations.

Both the cluster and structure analysis exhibited similar results distributing the whole population taken for this genetic diversity study into three groups. Though the overall result could not reveal a clear-cut means for the evolution of weedy rice prevailing in the state of Odisha, India, however, partly it paves the way of approach to understand the origin of weedy rice by further use of more number of molecular markers and domesticated genes.

Competing interests
The authors declare no competing interests.

Acknowledgements
The authors wish to thank Director, ICAR-National Rice Research Institute (formerly CRRI), Odisha for providing all the facilities required in the present studies.

Conflict of Interest: None declared

References

[1] Khush G. S. (2005) *Plant Molecular Biology,* 59, 1–6.
[2] Song B., Chuah T., Tam S. M. and Olsen K. M. (2014) *Molecular Ecology,* 23, 5003–5017.
[3] Gealy D. H., Agrama H. and Jia M. H. (2012) *Weed Science,* 60, 451–461.
[4] Hoagland R. E. and Paul R. V. (1978) *Weed Science*, 26619–625.
[5] Bres-Patry C., Lorieux M., Clément G., Bangratz M. and Ghesquière A. (2001) *Theoretical and Applied Genetics,* 102(1), 118–126.
[6] Mao-bai L., Hui W. and Li-ming C. (2015) *Rice Science,* 22(4), 180−188.
[7] Chung N. J. (2010) *Crop and Pasture Science,* 61, 911–917.
[8] Oka H.I. (1988) Origin of cultivated rice. Japanese Scientific Societies Press, Tokyo.
[9] Suh H. (2008) Weedy rice. Wild Crop Germplasm Center of Yeungnam University, Korea.
[10] Vaughan L. K., Ottis B. V., Prazak-Havey A. M., Bormans C. A., Sneller C., Chandler J. M. and Park W. D. (2001) *Weed Science,* 49, 468–476.
[11] Green J. M., Barker J. A., Marshall E. P., Froud-Williams R. J., Peters N. B., Arnold G. M., et al. (2001) *Molecular Ecology,* 10, 1035–1045.
[12] Prathepha, P. (2009) *Weed Biology and Management,* 9, 1-9.
[13] Pyšek P. and Prach K. (2003) *Biological Invasions,* 5(4), 337–348.
[14] Akasaka M., Ushiki J., Iwata H., Ishikawa R. and Ishii T. (2009) *Breeding Sci.* 59, 401–409.
[15] Reagon M., Thurber C. S., Gross B. L., Olsen K. M., Jia Y. L. and Caicedo A. L. (2010) *BMC Evolutionary Biology,* 10, 180.
[16] Grimm A., Fogliatto S., Nick P., Ferrero A. and Vidotto F. (2013) *Ecology and Evolution,* 3, 4786–4798.
[17] Thurber C.S., Jia M.H., Jia Y. and Caicedo A.L. (2012) M*olecular Ecology,* 22(3), 685-98.
[18] Londo J. P. and Schaal B. A. (2007) *Molecular Ecology,* 16, 4523-4535.
[19] Prathepha P. (2011) *AJCS,* 5(1), 49-54.
[20] Schaal, B. A. et al. (2010) Gene flow, biodiversity, and genetically modified crops: Weedy rice in Thailand. In: DeWoody, J. A. et al. eds. Molecular approaches in natural resource conservation and management. (Cambridge University Press, Cambridge).
[21] Gressel J. (2005) Crop ferality and volunteerism. (CRC Press, Boca Raton, 2005).
[22] Gross, B. L., Reagon M., Hsu S. C., Caicedo A. L., Jia Y. and Olsen K. M. (2010) *Molecular Ecology,* 19, 3380-3393.
[23] Zhang J., Lu Z., Dai W., Song X., Peng Y., Valverde B. E. and Qiang S. (2015) *Scientific Reports* 5.
[24] Doyle J. J. and Doyle J. L. (1990) *Focus* 12, 13-15.
[25] Rohlf F. J. (1998) NTSYSpc: numerical taxonomy and multivariate analysis system, version 202. Setauket, New York: Exeter Software.
[26] Liu K. and Muse S. V. (2005) *Bioinformatics* 21, 2128–2129.
[27] Peakall R. and Smouse P. E. (2012) *Bioinformatics,* 28, 2537-2539.
[28] Pritchard J.K., Stephens M. and Donnelly P. (2000) *Genetics,* 155, 945–959.
[29] Evanno G., Regnaut S. and Goudet J. (2005) *Molecular Ecology,* 14, 2611–2620.
[30] Earl D. A. and VonHoldt B. M. (2012) *Conservation Genetics Resources,* 4(2), 359-361.
[31] Zhang L., Dai W., Wu C., Song X., and Qiang S. (2012) *Genetic Resources and Crop Evolution,* 59, 399–410.
[32] Courtois B., Frouin J., Greco R., Bruschi G., Droc G., Hamelin C., Ruiz M., Clément G., Evrard J. C. and Van S. (2012) *Crop Science,* 52(4), 1663–1675.
[33] Jin L., Lu Y., Xiao P., Sun M., Corke H. and Bao J. (2010) *Theoretical and Applied Genetics,* 121(3), 475–487.
[34] Nachimuthu V. V., Muthurajan R., Duraialaguraja S., Sivakami R., Pandian B. A., Ponniah G., Gunasekaran, Swaminathan M., Suji K. K. and Sabariappan R. (2015) *Rice,* 8, 30.
[35] Neel J. V. (1973) "Private" genetic variants and the frequency of mutation among South American Indians. *Proceeding of National Academy of Science USA,* 70, 3311–3315.
[36] Slatkin M. (1985) *Evolution,* 39, 53–65.
[37] Jiang Z., Xia H., Basso B. and Lu B. (2012) *Theoretical and Applied Genetics,* 124, 309–322.
[38] Song Z., Wang Z., Feng Y., Yao N., Ji Y. and Lu B. (2015) *Journal of Systematics and evolution,* 9999, 1-9.
[39] Agrama H. and Eizenga G. (2008) *Euphytica,* 160(3), 339–355.
[40] Cao Q., Lu B. R., Xia H., Rong J., Sala F., Spada A. and Grassi F. (2006) *Annal of Botany,* 98, 1241-1252.
[41] Zhang J., Burgos N. R., Kun M. A., Zhou Y. J., Geng R. M. and Yu L. Q. (2008) *Rice Science,* 15(4), 295-302.
[42] Shao, J., Dai, W. M., Zhang, L. J., Song, X. L. and Qiang, S. (2011) *Acta Agronomica Sinica,* 37, 1324-1332.
[43] Xia H.B., Wang W., Xia H., Zhao W. and Lu B. R. (2011) *PLoS One,* 6(1), e16189.
[44] Horst L. and Wenzel G. (2007) Molecular marker systems in plant breeding and crop improvement (Vol. 55). Springer Science & Business Media.
[45] Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S. and Rafalski A. (1996) *Molecular Breeding,* 2(3), 225–238.
[46] Varshney R. K., Chabane K., Hendre P. S., Aggarwal R. K. and Graner A. (2007) *Plant Science,* 173(6), 638-649.
[47] Liakat A. M., McClung A. M., Jia M. H., Kimball J. A., McCouch S. R., Georgia C. E. (2011) *Crop Science*, 51(5), 2021–2035.
[48] Pusadee T., Schaal B., Rerkasem B. and Jamjod S. (2013) *Genetic Resources and Crop Evolution,* 60, 335-353.
[49] Kioko W.F., Musyoki M.A., Piero N.M., Muriira K.G., Wavinya N.D., Rose L., Felix M. and Ngithi N.L. (2015) *Journal of Phylogenetics and Evolutionary Biology,* 3, 158.
[50] Shahidullah S.M., Hanafi M.M., Ashrafuzzaman M., Ismail M. R. and Khair A. (2009) *African Journal of Biotechnology,* 8, 1238-1246.