View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by AJOL - African Journals Online

African Journal of Biotechnology Vol. 10(35), pp. 6675-6685, 13 July, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.2279
ISSN 1684–5315 © 2011 Academic Journals

Full Length Research Paper

Suitability of a selected set of simple sequence repeats (SSR) markers for multiplexing and rapid molecular characterization of African rice (*Oryza glaberrima* Steud.)

Khady Nani Dramé[1]*, Ines Sanchez[1], Glenn Gregorio[2] and Marie Noëlle Ndjiondjop[1]

[1]Africa Rice Center 01 BP 2031 Cotonou, Benin.
[2]International Rice Research Institute (IRRI) DAPO Box 7777 Metro Manila, Philippines.

Accepted 5 May, 2011

African rice (*Oryza glaberrima*) was neglected for a long time by researchers but gained attention with the success of NERICA varieties. While AfricaRice holds a collection of about 2,500 *O. glaberrima* samples only four accessions have been used as NERICA parents. This collection needs to be characterized using rice descriptors and molecular markers to fully exploit the whole diversity of this species. To allow a quick and accurate molecular characterization of African rice, a set of 30 SSR markers were identified and tested on a subset of 74 African rice samples. The results showed that these markers were highly polymorphic, could be multiplexed in 8 panels without overlapping alleles and were able to distinguish the genetic groups of African rice; the two major ecotypes and intermediates between *Oryza sativa* and *O. glaberrima* species.

Key words: Rice, *Oryza glaberrima*, simple sequence repeats (SSR) markers, genetic diversity, NERICA (New Rice for Africa).

INTRODUCTION

Only two species of cultivated rice exist in the world: *Oryza glaberrima* or African rice and *Oryza sativa* or Asian rice. Native to sub-Saharan Africa (SSA), *O. glaberrima* is thought to have been domesticated from the wild ancestor *Oryza barthii* some 2,000 to 3,000 years ago in the inland delta of the upper Niger river in what is today Mali, long before the introduction of *O. sativa* varieties by the Portuguese (Linares, 2002). Because of its low yield potential due to high shattering and lodging susceptibility, *O. glaberrima* is being increasingly replaced by the higher yielding *O. sativa* varieties (Linares, 2002). However, in margin areas and under poor crop management or adverse ecological conditions, *O. glaberrima* is still favored by farmers for its adaptability and resistance to multiple constraints (Jones et al., 1997; Linares, 2002; Sarla and Mallikarjuna, 2005).

For years, *O. glaberrima* was of little interest for research but the success of the NERICA (New Rice for Africa) varieties developed by AfricaRice reversed the trend. These varieties widely grown in SSA are the first interspecific *O. sativa* x *O. glaberrima* cultivars. They combined the hardiness of *O. glaberrima* and the yield potential of *O. sativa* (Jones et al., 1997). However, there are still gaps between the current NERICA varieties and the *O. glaberrima* varieties regarding the resistance to local constraints of SSA (Futakuchi and Sie, 2009). Recent studies showed that there is still a tremendous amount of unexploited genetic diversity in the primary gene pool of African rice that can be used to enhance the diversity in local germplasms and improve their performance under diverse agroecological conditions (Futakuchi and Sie, 2009; Mohapatra, 2010; Thiemele et al., 2010). The Genbank of Africa Rice holds a collection

*Corresponding author. E-mail: k.drame@cgiar.org. Tel: +22921350188. Fax: +22921350556.

of about 2,500 samples of African rice which still need to be fully characterized for accurate referencing and identification of new resistance sources that could be used in breeding programs. Unlike *O. sativa* species which is classified into, *indica* and *japonica*, no subspecies can be recognized for *O. glaberrima* species (Second, 1982). However, the existence of two major ecotypes is admitted, a floating photosensitive type grown in deep water, including coastal mangrove areas and an early erect ecotype grown in upland or moderately inundated lowlands (Ghesquiere et al., 1997). Using simple sequence repeats (SSR) markers Semon et al. (2005) revealed the population structure of African Rice. Five distinct genetic groups were identified out of which two clustered with *O. sativa* subspecies *indica* and *japonica*. In order to fully exploit the genetic diversity of this species, the whole collection of 2,500 samples should be characterized and each sample assigned to the right genetic group. For this purpose, the availability of molecular markers that can distinguish the different genetic groups of African rice at a reduced cost and time will be of great interest. Therefore, a set of 30 SSR markers has been identified and tested on a subset of 74 samples to perform quick and accurate genotyping. Specifically, we tested the possibility of multiplexing these markers to quicken the genotyping and we assessed their informativeness and efficiency to distinguish the different genetic groups of African rice.

MATERIALS AND METHODS

Plant material

A set of 82 samples were used in the study comprising 74 *O. glaberrima* samples, six *O. sativa* and two wild species (Table 1). The *O. glaberrima* and *O. sativa* samples were provided by the AfricaRice Genbank and the wild species were obtained through collaborators. The two sativa subspecies were represented by three popular varieties each. The two wild species, one *Oryza barthii* and one *Oryza longistaminata* constituted the control samples as well as the *O. sativa* varieties. The *O. glaberrima* samples originated from eight different countries and are adapted to various ecosystems (Table 1). For a greater purity, seeds were collected from a single panicle and pregerminated on humidified filter papers in Petri dishes. The dormancy of *O. glaberrima* samples was broken prior to pregermination by removing the seed coat and cutting the opposite edge of the embryo. Three plantlets per sample were then grown in pots.

DNA extraction

Young leaves from 1 month-old plant were bulk harvested and frozen at -20° until DNA extraction. Leaf tissue of about 100 mg was weighed and transferred inside 12 x 8 well-racks containing two stainless steel grinding balls (OPS diagnostics). One more ball was put on top of the samples to ensure proper grinding of the leaf tissues. The samples were processed in a GenoGrinder 2000 (Spex CertiPrep Inc, USA), following the manufacturer's instructions, at 1500 strokes per minute for 3 to 5 min. Genomic DNA was isolated from the grinded samples using miniprep DNA extraction method modified from Risterucci et al. (2000). The DNA dry pellets were dissolved in 30 µl of TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8). Aliquots of 2 µl of freshly extracted DNA were electrophoresed on 1% agarose gel, stained with ethidium bromide and visualized under an UV transilluminator (Alpha Imager®HP, AlphaInnotech) to assess the quality of the DNA extracted. The concentration of the samples was determined using a Nanodrop D-1000 spectrophotometer. They were subsequently diluted to 5 ng/µl.

Genotyping

The diluted DNA samples were subjected to PCR amplification using 30 simple sequence repeat (SSR) markers well dispersed on the rice genetic map (Table 2). The markers were selected on the basis of the polymorphism they had revealed in rice in our previous work. The sequences of the markers were available on Gramene website (www.gramene.org/microsat/). The forward primer for each of the marker was labeled at the 5' end of the oligonucleotide with one of the following fluorescent dyes: 6-FAM™ (Blue), VIC™ (Green), NED™ (Yellow), PET™ (Red) (Applied Biosystems Inc., USA). PCR amplifications were performed separately for each marker using 96-well plates. The total reaction volume of 10 µl consisted of 5 ng DNA, 1× PCR buffer (Roche), 2.5 µmol each of the forward and reverse primers, 2 mmol dNTPs mix and 0.5 unit *Taq* DNA polymerase (Roche). Amplifications were carried out in a gradient thermal cycler (BioRad) using the following program: initial denaturing at 94° for 5 min followed by 35 cycles of 94° for 30 s, 55° annealing for 30 s and 72° for 30 s and a further primer extension at 72° for 2 min. Prior to visualization by an automated capillary electrophoresis, 6 PCR products per marker were randomly picked and migrated on agarose gel 2% to check the amplification.

Capillary electrophoresis

Based on the allele size published on Gramene website, eight panels were composed using non overlapping markers labeled with different dyes (Table 2). For each dye color, appropriate dilution of PCR products was made prior to multiplexing (3 to 4 different markers). A master mix of 10 µl consisting of 1 µl each diluted PCR product, 5.7 or 6.7 µl formamide (Hi-Di) and 0.3 µl GeneScan-500 LIZ size standard (Applied Biosystems Inc., USA) was prepared into optical 96-well MicroAmp plates (Applied Biosystems Inc., USA) for DNA fragments separation during electrophoresis. The master mix was then denatured at 94° for 5 min. Fragment separation was performed using the ABI PRISM 3100 genetic analyzer machine containing POP 4 (performance optimized polymer) matrix and urea following the protocol developed at the Central Biotechnology Laboratory of IITA, Ibadan, Nigeria.

Data collection and analysis

The data obtained were analyzed using GeneMapper software version 4.0 (Applied Biosystems Inc., USA). A single or pair of peaks was detected as expected for the co-dominant markers such as SSRs. Fragment sizes were automatically calculated with reference to the internal lane size standard GeneScan-500 LIZ ranging from 35 to 500 bp. Because of their proximity with primer peaks, the size marker peaks corresponding to 35 and 50 bp were excluded from the analysis. A program developed by Ghosh et al. (1997) was used to increase the precision of allele sizing and to reduce genotyping error rates. This program created, for each allele

Table 1. Rice genotypes used in the study, species, country of origin and ecosystem as referenced in AfricaRice Genbank database.

Number	Genotype	Species	Origin	Ecosystem
1	IRGC101937	*O. barthii*	Senegal	Unknown
2	Unknown	*O. longistaminata*	Mali	Unknown
3	CG 14	*O. glaberrima*	Senegal	Upland
4	CG 17	*O. glaberrima*	Senegal	Upland
5	DC KONO	*O. glaberrima*	Sierra Leone	Upland
6	Gbobye	*O. glaberrima*	Guinea	Upland
7	IG 02	*O. glaberrima*	Cote d'Ivoire	Upland
8	Pa DC KONO	*O. glaberrima*	Sierra Leone	Upland
9	Saliforeh	*O. glaberrima*	Sierra Leone	Upland
10	Salikatato	*O. glaberrima*	Guinea	Upland
11	TOG 5282	*O. glaberrima*	Cote d'Ivoire	Upland
12	TOG 5286	*O. glaberrima*	Nigeria	Deep Forest swamp
13	TOG 5287	*O. glaberrima*	Nigeria	Deep Forest swamp
14	TOG 5290	*O. glaberrima*	Nigeria	Deep Forest swamp
15	TOG 5307	*O. glaberrima*	Nigeria	Upland
16	TOG 5314	*O. glaberrima*	Nigeria	Rainfed lowland
17	TOG 5324	*O. glaberrima*	Nigeria	Shallow forest swamp
18	TOG 5326	*O. glaberrima*	Nigeria	Shallow forest swamp
19	TOG 5378	*O. glaberrima*	Nigeria	Deep forest swamp
20	TOG 5390	*O. glaberrima*	Nigeria	Floating rice
21	TOG 5400	*O. glaberrima*	Nigeria	Floating rice
22	TOG 5404	*O. glaberrima*	Nigeria	Deep forest swamp
23	TOG 5420	*O. glaberrima*	Nigeria	Shallow forest swamp
24	TOG 5423	*O. glaberrima*	Nigeria	Shallow forest swamp
25	TOG 5429	*O. glaberrima*	Nigeria	Shallow forest swamp
26	TOG 5437	*O. glaberrima*	Nigeria	Shallow forest swamp
27	TOG 5439	*O. glaberrima*	Nigeria	Deep forest swamp
28	TOG 5447	*O. glaberrima*	Nigeria	Shallow forest swamp
29	TOG 5453	*O. glaberrima*	Nigeria	Shallow forest swamp
30	TOG 5458	*O. glaberrima*	Nigeria	Shallow forest swamp
31	TOG 5464	*O. glaberrima*	Nigeria	Shallow forest swamp
32	TOG 5473	*O. glaberrima*	Nigeria	Floating rice
33	TOG 5486	*O. glaberrima*	Nigeria	Shallow forest swamp
34	TOG 5491	*O. glaberrima*	Nigeria	Shallow forest swamp
35	TOG 5500	*O. glaberrima*	Nigeria	Shallow forest swamp
36	TOG 5514	*O. glaberrima*	Nigeria	Shallow forest swamp
37	TOG 5523	*O. glaberrima*	Nigeria	Shallow forest swamp
38	TOG 5533	*O. glaberrima*	Nigeria	Upland
39	TOG 5556	*O. glaberrima*	Nigeria	Shallow forest swamp
40	TOG 5566	*O. glaberrima*	Nigeria	Shallow forest swamp
41	TOG 5591	*O. glaberrima*	Ghana	Shallow forest swamp
42	TOG 5602	*O. glaberrima*	Ghana	Shallow forest swamp
43	TOG 5639	*O. glaberrima*	Nigeria	Upland
44	TOG 5650	*O. glaberrima*	Nigeria	Rainfed lowland
45	TOG 5666	*O. glaberrima*	Nigeria	Shallow forest swamp
46	TOG 5672	*O. glaberrima*	Nigeria	Shallow forest swamp
47	TOG 5674	*O. glaberrima*	Nigeria	Irrigated lowland
48	TOG 5675	*O. glaberrima*	Nigeria	Irrigated lowland
49	TOG 5681	*O. glaberrima*	Nigeria	Deep forest swamp

Table 1. Contd.

50	TOG 5687	*O. glaberrima*	Nigeria	Shallow forest swamp
51	TOG 5695	*O. glaberrima*	Nigeria	Rainfed lowland
52	TOG 5775	*O. glaberrima*	Liberia	Upland
53	TOG 5803	*O. glaberrima*	Liberia	Upland
54	TOG 5820	*O. glaberrima*	Liberia	Upland
55	TOG 5885	*O. glaberrima*	Liberia	Upland
56	TOG 5919	*O. glaberrima*	Liberia	Upland
57	TOG 5923	*O. glaberrima*	Liberia	Upland
58	TOG 5953	*O. glaberrima*	Nigeria	Upland
59	TOG 5963	*O. glaberrima*	Nigeria	Rainfed lowland
60	TOG 5969	*O. glaberrima*	Nigeria	Upland
61	TOG 5989	*O. glaberrima*	Nigeria	Rainfed lowland
62	TOG 6000	*O. glaberrima*	Nigeria	Upland
63	TOG 6007	*O. glaberrima*	Nigeria	Upland
64	TOG 6038	*O. glaberrima*	Nigeria	Rainfed lowland
65	TOG 6181	*O. glaberrima*	Nigeria	Rainfed lowland
66	TOG 6211	*O. glaberrima*	Nigeria	Unknown
67	TOG 6308	*O. glaberrima*	Liberia	Upland
68	TOG 6334	*O. glaberrima*	Liberia	Upland
69	TOG 6356	*O. glaberrima*	Liberia	Upland
70	TOG 6679	*O. glaberrima*	Liberia	Upland
71	TOG 6710	*O. glaberrima*	Liberia	Upland
72	TOG 6767	*O. glaberrima*	Liberia	Upland
73	TOG 7106	*O. glaberrima*	Mali	Shallow forest swamp
74	TOG 7206	*O. glaberrima*	Cote d'Ivoire	Upland
75	TOG 7345	*O. glaberrima*	Nigeria	Shallow forest swamp
76	TOG 7420	*O. glaberrima*	Sierra Leone	Upland
77	BG90-2	*O. sativa indica*	Sri Lanka	Irrigated lowland
78	Bouake 189	*O. sativa indica*	Cote d'Ivoire	Irrigated lowland
79	IR 64	*O. sativa indica*	Philippines	Lowland
80	Azucena	*O. sativa japonica*	Philippines	Upland
81	IAC 165	*O. sativa japonica*	South America	Upland
82	Moroberekan	*O. sativa japonica*	Africa	Upland

of a marker, bins that are mean size in base pairs of alleles for that particular allele category, rounded off to the nearest whole number. Adjusted bins were exported in GeneMapper V4.0 and reanalyzed to correct the allele sizing. Statistical parameters defining the diversity such as allele number, frequency of major allele, gene diversity, heterozygosis and polymorphism information content (PIC) were determined using PowerMarker V3.25. Population structure of the samples was inferred using a model-based program, STRUCTURE V2.1. The following parameters: maximum populations (K = 10), length of Burnin period equal to 50,000, number of Markov chain Monte Carlo (MCMC) repetitions after Burnin equal to 200,000 and 10 independent runs for each K were fixed for the simulation and the ancestry model used was Admixture model with the limit to assign a sample to a population fixed at 75% genome ancestry. STRUCTURE may tend to identify more populations than are biologically relevant (Falush et al., 2003); thus, the true number of populations (K) was calculated according to Evanno et al. (2005). Among the 10 runs of the true K, the one with the highest log likelihood (posterior probability) was used to identify individuals assigned to populations and admixtures in STRUCTURE V2.1. The genetic distance between genotypes was calculated with DARwin V5.0.158 using simple matching method. The phylogenetic tree was built using the unweighted neighbor-joining (UWNJ) algorithm of DARwin V5.0.158.

RESULTS

Allelic diversity detected with fluorescent SSR markers

The SSR markers used in this study allowed clear amplification of African rice DNA samples. 21 of them were included in the Panel of 50 standard SSR markers used by the generation challenge program (GCP) for rice diversity analyses. Based on the allele size range

Table 2. SSR markers selected, location on chromosome (Chr.) expressed in centimorgan (cM), range of allele size (www.gramene.org) and dye color.

Panel	Marker	Chr.	Position_cM	Allele size (bp)	Dye color
1	RM125	7	24.8	105-147	Red
	RM11	7	47	118-151	Yellow
	RM507	5	0	234-257	Blue
2	RM1	1	29.7	67-119	Red
	RM455	7	65.7	127-144	Yellow
	RM154	2	4.8	148-230	Blue
	RM431	1	178.3	233-261	Green
3	RM287	11	68.6	82-118	Red
	RM261	4	35.4	119-140	Yellow
	RM333	10	110.4	150-270	Blue
	RM171	10	73	307-347	Green
4	RM312	1	71.6	86-106	Red
	RM408	8	0	112-128	Yellow
	RM234	7	88.2	133-163	Blue
	RM7	3	64	167-183	Green
5	RM109	2	4.8	86-105	Red
	RM3	6	74.9	116-142	Yellow
	RM338	3	108.4	178-184	Blue
	RM219	9	11.7	190-232	Green
6	RM447	8	124.6	95-146	Red
	RM223	8	80.5	133-164	Yellow
	RM19	12	20.9	192-250	Blue
	RM124	4	150.1	257-289	Green
7	RM277	12	57.2	104-121	Red
	RM152	8	9.4	133-157	Yellow
	RM316	9	1.8	194-216	Blue
	RM536	11	55.1	223-247	Green
8	RM237	1	115.2	105-153	Red
	RM249	5	65.8	105-171	Yellow
	RM433	8	116	216-248	Blue

reported in Gramene for reference sativa varieties, eight panels were constituted combining each from three to four different markers. Well separated peaks were obtained for each combined markers of a given panel and for a given genotype (Figure 1). All markers were polymorphic detecting 3 to 19 alleles per locus with an average of 8.4 alleles/locus (Table 3). The markers RM333 and RM249 detected the maximum number of alleles. The frequency of the major allele ranged from 0.2 to 0.88 with an average of 0.59. A total of 252 alleles were detected using the whole set of 30 markers. The database of allelic frequencies showed that rare alleles (with a frequency < 0.05) accounted for 61.1% of all alleles, while intermediate (0.05 < frequency < 0.30) and abundant alleles (frequency > 0.30) accounted respectively for 25.8 and 13.1% of all the detected alleles. Allele size ranged from 81 to 347 bp. Mean PIC value was 0.52 ranging from 0.2 to 0.9. At 6 loci (RM1, RM431, RM234, RM7, RM19 and RM124) out of the 30, all samples studied were homozygous. At the remaining

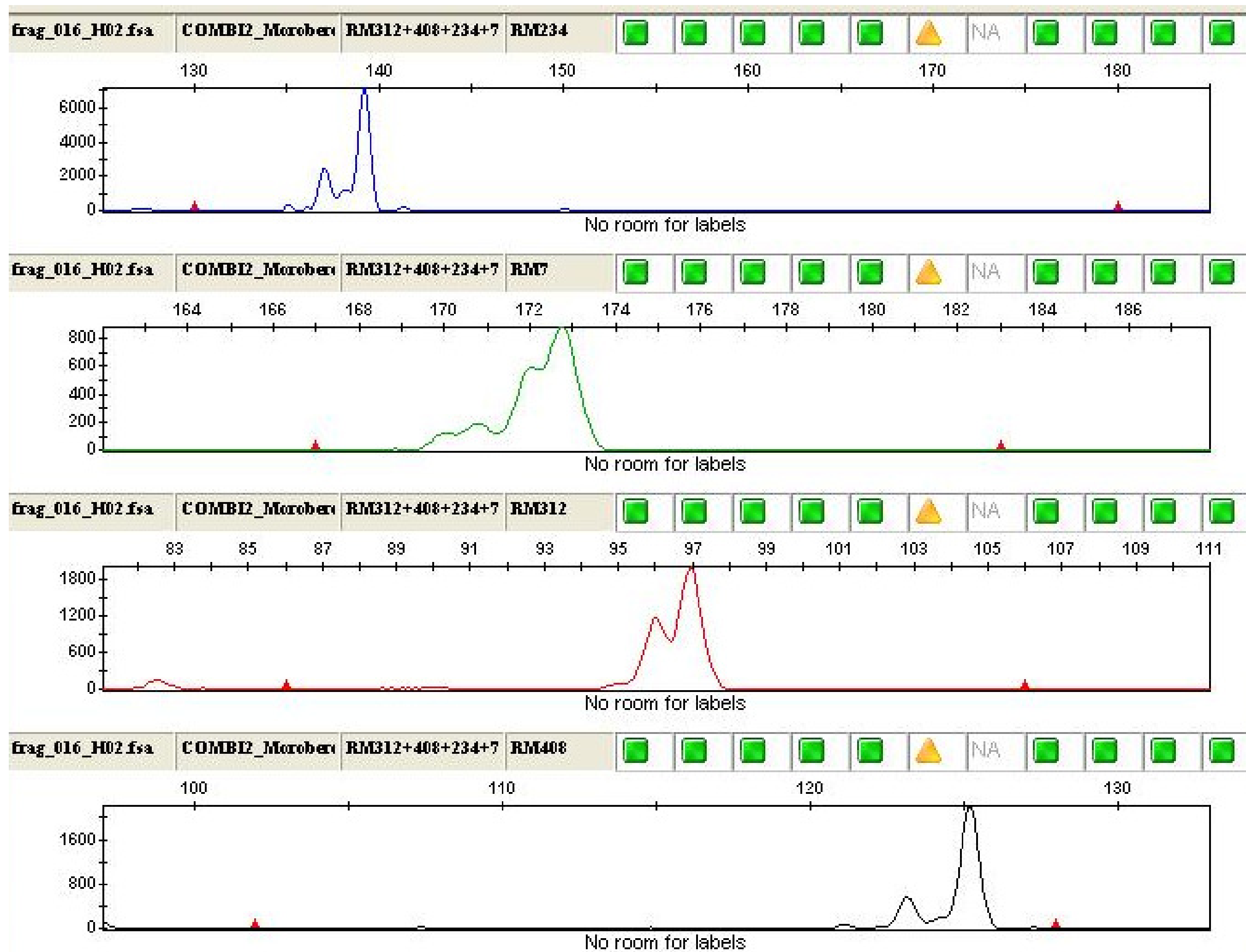

Figure 1. Electrophoregram obtained for one of the samples using panel 4 which is a combination of RM312, RM408, RM234 and RM7.

loci, one or more heterozygous individuals were detected, with RM125 showing the maximum of heterozygous (Table 3).

Global diversity structure and genetic relationships

Genetic structure and diversity were determined using ancestry-based grouping and phylogenetic representation. Using the method described by Evanno et al. (2005) we found the modal value of the distribution of ΔK (a parameter derived from the number of populations K) to be located at K equal 3 which suggested the existence of three distinct populations (Pop1, Pop2 and Pop3) in our *O. glaberrima* samples. Since the limit to assign a sample to a population was fixed at 75% genome ancestry, we found that there were a few samples that shared ancestry between the different populations identified. The samples TOG7106, TOG5458, TOG5491 and TOG6000 shared ancestry with Pop2 and Pop3. The proportion of genome ancestry shared was on average 33.8 and 66.1%. The sample TOG5486 shared ancestry with Pop1 and Pop3 (64.9 and 34.8% ancestry shared) and TOG5404 with all three populations (17.7, 63.2 and 19.4% respectively with Pop1, Pop2 and Pop3). The remaining samples were assigned to the three populations as followed: 5 samples in Pop1, 42 samples in Pop2 and 21 samples in Pop3.

UWNJ tree cluster analysis enabled the identification of 4 groups within the *O. glaberrima* samples (Figure 2). The group G1 represented *O. glaberrima* samples that clustered with the *O. sativa* species. The *O. glaberrima* samples Saliforeh, PaDcKono and DcKono clustered with the *japonica* subspecies, while TOG6007 and TOG5556

Table 3. Diversity detected within the subset of the 82 samples at 30 different loci.

Panel	Marker	Frequency	Missing (%)	N.A.	Gene Div.	He.	PIC	Allele size (bp)
1	RM125	0.49	0.00	7	0.60	0.86	0.51	113-136
	RM11	0.47	0.00	8	0.64	0.05	0.58	113-141
	RM507	0.86	9.76	3	0.26	0.02	0.24	248-256
2	RM1	0.85	0.00	5	0.29	0.00	0.28	83-108
	RM455	0.57	8.54	9	0.62	0.02	0.58	120-149
	RM154	0.78	0.00	10	0.40	0.02	0.39	164-195
	RM431	0.43	0.00	6	0.66	0.00	0.59	241-251
3	RM287	0.65	2.44	8	0.55	0.02	0.52	86-112
	RM261	0.39	1.22	6	0.70	0.02	0.64	116-133
	RM333	0.24	3.66	19	0.90	0.07	0.89	153-263
	RM171	0.52	7.32	8	0.61	0.02	0.53	321-347
4	RM312	0.50	0.00	10	0.69	0.02	0.65	93-110
	RM408	0.85	0.00	6	0.29	0.03	0.27	99-125
	RM234	0.22	0.00	12	0.87	0.00	0.86	121-163
	RM7	0.86	0.00	6	0.27	0.00	0.26	162-180
5	RM109	0.37	1.22	9	0.75	0.07	0.70	81-110
	RM3	0.58	21.96	7	0.63	0.05	0.59	104-142
	RM338	0.89	1.22	3	0.21	0.07	0.20	177-182
	RM219	0.21	3.66	17	0.91	0.06	0.90	193-233
6	RM447	0.45	0.00	7	0.63	0.02	0.55	97-127
	RM223	0.31	0.00	10	0.78	0.03	0.75	131-161
	RM19	0.82	0.00	6	0.33	0.00	0.32	203-247
	RM124	0.88	1.22	4	0.23	0.00	0.21	264-287
7	RM277	0.79	1.22	4	0.37	0.02	0.34	114-122
	RM152	0.81	1.22	10	0.34	0.09	0.34	114-152
	RM316	0.56	3.66	9	0.65	0.06	0.62	165-223
	RM536	0.79	10.98	7	0.37	0.02	0.36	194-233
8	RM237	0.69	0.00	8	0.52	0.02	0.50	121-135
	RM249	0.27	0.00	19	0.86	0.07	0.85	102-154
	RM433	0.69	0.00	9	0.51	0.02	0.47	219-307
	Mean	0.59	2.64	8.40	0.55	0.06	0.52	
	Min	0.21	0.00	3.00	0.21	0.00	0.20	
	Max	0.89	21.96	19.00	0.91	0.86	0.90	

Frequency; major allele frequency; N.A, number of alleles detected; He, proportion of heterozygous individuals in the samples at a given loci; gene Div., gene diversity; PIC, polymorphism information content.

are grouped with the *indica* subspecies. These *O. glaberrima* samples, genetically close to *O. sativa* species were the same than the ones identified with the ancestry-based method in Pop1. The other groups, G2 and G3 numbered respectively 21 and 23 samples exclusively *O. glaberrima* species and the group G4 included the *O. barthii* sample among the 24 samples. The sample TOG5404 clustered in none of the four groups as shown with the first method. The genetic groups G2 and G3 corresponded to the population Pop2 determined with STRUCTURE while the group G4 corresponded to the population Pop3. Only one sample of G4, TOG6181 was assigned to Pop2 instead of Pop3.

No clear distribution of these genetic groups based on the country of origin of the samples could be observed. But considering the ecosystem of adaptation, we noted that majority of G4 individuals are upland samples (65.22%), while majority of G2 (66.67%) and G3 (52.17%) were samples adapted to swamp ecosystem (Table 4). Lowland-adapted samples were present only in G3 and G4. Samples adapted to rainfed lowland were present in G4, while those adapted to irrigated lowland were present in G3 (Tables 1 and 4). The lowland NERICA parents TOG5674, TOG5675 and TOG5681 were assigned to the group G3, while the upland NERICA

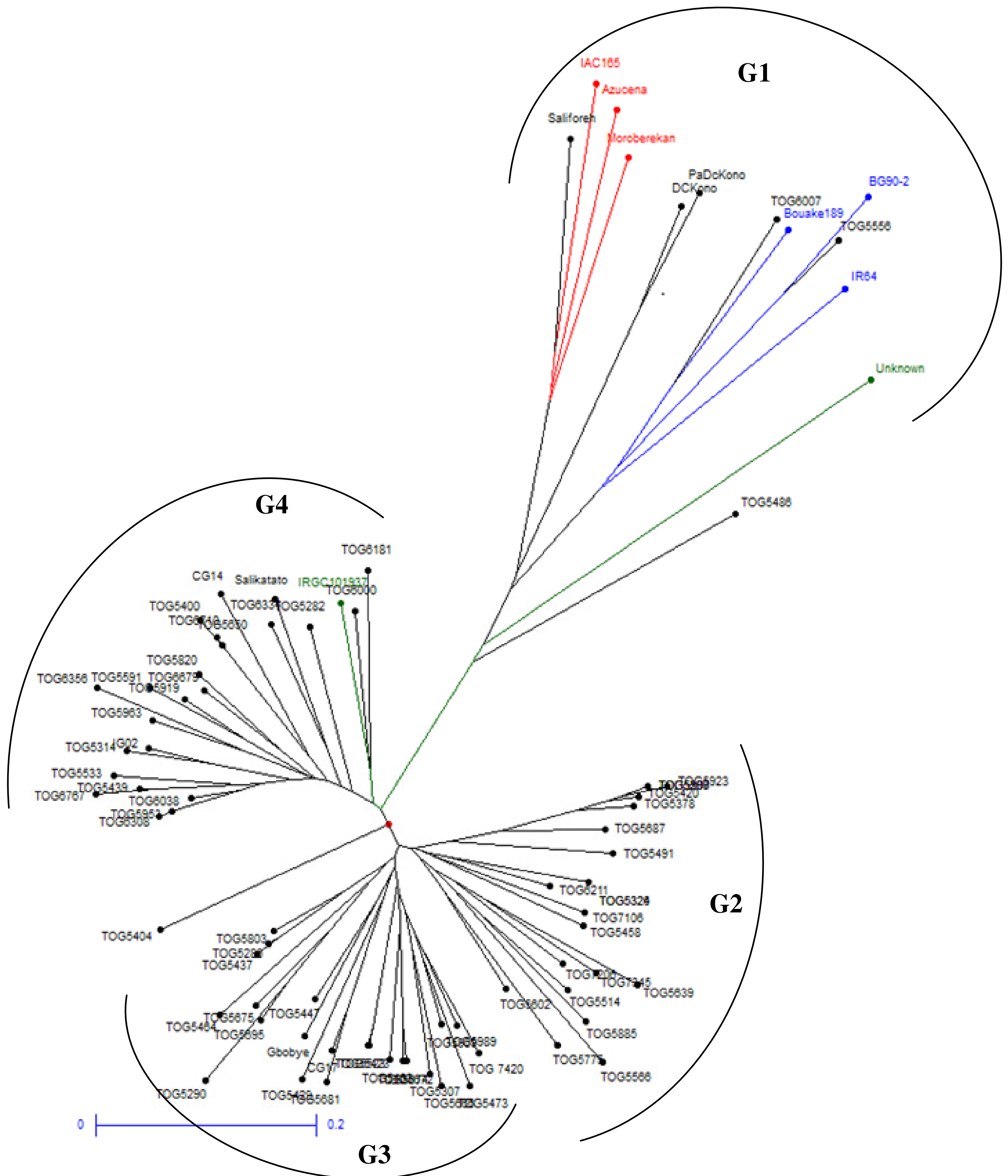

Figure 2. Genetic groups revealed by the UWNJ method. Wild species are written in green, *O. sativa indica* in blue and *O. sativa japonica* in red. Plain circles delimited the phylogenetic groups G1, G2, G3 and G4.

Table 4. Distribution of the 4 genetic groups identified by UNWNJ analysis in the ecosystems of adaptation of the *O. glaberrima* samples. The number of samples in each genetic group is followed by its percentage out of the total number of individuals in each group which is respectively 6, 21, 23 and 23 for G1, G2, G3 and G4.

Ecosystem	Genetic group identified					
	G1	G2	G3	G4	NA	Total
Floating	0 (0)	1 (4.76)	1 (4.35)	1 (4.35)	0	3
Upland	4 (66.67)	5 (23.8)	6 (26.09)	15 (65.22)	0	30
Lowland*	0 (0)	0 (0)	4 (17.39)	5 (21.74)	0	9
Swamp**	2 (33.33)	14 (66.67)	12 (52.17)	2 (8.69)	1 (100)	31
Unknown	0 (0)	1 (4.76)	0 (0)	0 (0)	0	1

*Both irrigated and rainfed lowland;** both shallow forest and deep forest swamp;NA, assigned to none of the four groups.

parent CG14 was present in the group G4.

DISCUSSION

In this work, 30 fluorescent SSR markers were tested on 82 samples including 74 samples of *O. glaberrima* and eight controls (6 *O. sativa* and 2 wild species) to assess their suitability for multiplexing, their informativeness and their efficiency to distinguish the different genetic groups of African rice. These SSR markers primarily designed from *O. sativa* (www.gramene.org, McCouch et al., 2002) were easily amplified in *O. glaberrima* and exhibited several alleles. The multiplexing into eight panels and the use of fluorescently labeled SSR markers in semi automated fashion greatly contributed to quicken the data acquisition and accuracy. We found a high level of genetic variability throughout the genome using this set of well-distributed markers. All the panels were polymorphic detecting 3 to 19 alleles per locus (8.4 on average), a total of 252 alleles and mean PIC of 0.52 over the 82 samples. Within the 74 *O. glaberrima* samples, the average number of alleles per locus obtained was 7.2 and mean PIC was 0.45. Compared with the other studies on African rice, Semon et al. (2005) and Barry et al. (2007) who respectively worked on 198 *O. glaberrima* accessions from 12 countries using 93 SSR markers and 26 *O. glaberrima* accessions from Maritime Guinea with 11 SSRs, the number of alleles per locus obtained in this study (NA = 7.2) was higher than that obtained by Barry et al. (NA = 4), but lower than what was obtained by Semon et al. (NA = 9.4). These differences were likely to be due to the sampling. Barry et al. (2007) worked on a lower number of *O. glaberrima* samples originating from one country, while Semon et al. (2005) worked on a higher number of samples originating from more countries than us. However, the PIC of 0.45 obtained in this study was the same as that obtained by Barry et al. (2007), while it was higher than what was obtained by Semon et al. (2005) (PIC= 0.34). This could be due to the fact that the SSR markers used in this study were selected for their high polymorphism in African rice.

Independently of the method used, there were the same *O. glaberrima* samples that clustered with *O. sativa* subspecies (*indica* and *japonica*). Some of them clustered with *O. sativa* subspecies *indica* and others with the *japonica* subspecies as shown with the phylogenetic tree. Such a close genetic relationship between *O. sativa* and some *O. glaberrima* has been reported in other studies (Semon et al., 2005; Barry et al., 2007). The long cohabitation of the two cultivated species over hundreds of years in West Africa may have induced an introgression of *O. sativa* into *O. glaberrima* which resulted in the creation of intermediates (Semon et al., 2005; Nuijten et al., 2010). The remaining *O. glaberrima* samples clustered into three groups (G2, G3 and G4) where G2/G3 was constituted with a majority of samples adapted to swamp (mangrove) areas and G4 with a majority of upland samples. These groups could be representative of the two major ecotypes described by Ghesquière et al. (1997). These findings were also consistent with the population structure determined by Semon et al. (2005), where two of the genetic groups clustered with indica and japonica subspecies and the other three reflected the ecological adaptation to different environment. However, the proportion of ancestry shared with *O. sativa* species that we found for PaDcDono, DcKono and Saliforeh (99.8%) was much higher than the 54% reported by Semon et al. (2005). This resulted probably from the lower number of markers that we used.

The use of African rice in rice improvement resulted in the creation of NERICA varieties adapted to upland and lowland. The proportion of donor genome in these interspecific progenies was estimated using SSR markers. In the upland NERICA and sister lines, the proportion of *O. glaberrima* genome was estimated to account for 6.3% (Semagn et al., 2007), while in lowland NERICA and sister lines, the average contribution of donor parent at different backcross generations was 7.9% (Ndjiondjop et al., 2008). This low proportion of donor parent may explain the existing gap between current NERICAs and their *O. glaberrima* parents

regarding morphological characters and resistance to environmental constrains like weeds competitiveness and resistance to local diseases (Futakuchi and Sie, 2009). Both *O. sativa* and *O. glaberrima* are AA genome species, but the F_1 plants derived from crosses between the two species are generally highly pollen sterile (Ghesquiere et al., 1997). This substantially reduced the use of *O. glaberrima* genome in interspecific breeding. Higher compatibility in interspecific hybridization could be obtained by using *O. glaberrima* genotypes that are genetically close to *O. sativa*. Crosses between those glaberrimas and elite *O. sativa* varieties could result in more fertile F_1 lines. However, several authors (Ariyo 1987; Peeters and Martinelli, 1989; Souza and Sorrells, 1991) pointed out that crosses designed between genetically distant genotypes should produce higher variances in segregating populations than crosses between related genotypes. Individuals from different genetic groups may be the best candidates to serve as parents for breeding if sterility in the F_1 progenies can be overcome. This will now be possible with the interspecific bridges created between the two species (Semon, personal communication).

Conclusion

In sub-Saharan Africa, where the majority of the rice farmers are resource-poor, the use of natural genetic diversity to develop good performing varieties well adapted to their specific environmental and cropping conditions may be the best option to increase rice productivity and alleviate poverty. The choice of donor parents in a breeding scheme should then be based on the knowledge of the genetic group to which these parents belong. The information generated in this study provided a reliable set of markers that can be used to quickly assign in the right genetic group each of the 2,500 samples of African rice stored in the Genbank of AfricaRice. This combined with information collected during their agro morphological characterization will give a clear understanding of the genetic diversity of the entire collection.

ACKNOWLEDGEMENTS

The authors are grateful to Maria Kolesnikova-Allen for her training in the use of DARwin and GeneMapper software. This work was funded through a GCP project (GCP SP3 Project G4007.08) and K. N. Dramé was financially supported by an L'OREAL/UNESCO fellowship.

REFERENCES

Ariyo OJ (1987). Multivariate analysis and the choice of parents for hybridization in Okra (*Abelmoschus esculentus* (L.) Moench). Theor. Appl. Genet. 74: 361-363.

Barry MB, Pham JL, Noyer JL, Billot C, Courtois B, Ahmadi N (2007). Genetic diversity of the two cultivated rice species (*O. sativa* & *O. glaberima*) in Maritime Guinea. Evidence for interspecific recombination. Euphytica, 154: 127-137.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611-2620.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 164: 1567-1587.

Futakuchi K, Sie M (2009). Better exploitation of African Rice (*Oryza glaberrima* Steud.) in varietal development for resource-poor farmers in West and Central Africa. Agric. J. 4 (92): 96-102.

Ghesquiere A, Sequier J, Second G, Lorieux M (1997). First steps towards a rational use of African rice, *Oryza glaberrima*, in rice breeding through a 'contig line' concept. Euphytica, 96: 31-39.

Ghosh S, Karanjawala ZE, Hauser ER, Ally D, Knapp JI, Rayman JB, Musick A, Tannenbaum J, Te C, Shapiro S, Eldridge W, Musick T, Martin C, Smith JR, Carpten JD, Brownstein MJ, Powell JI, Whiten R, Chines P, Nylund SJ, Magnuson VL, Boehnke M, Collins FS (1997). Methods for precise sizing, automated binning of alleles, and reduction of error rates in large-scale genotyping using fluorescently labeled dinucleotide markers. FUSION (Finland-U.S. Investigation of NIDDM Genetics) Study Group. Genome Res. 7: 165-178.

Jones MP, Dingkuhn M, Aluko GK, Semon M (1997). Interspecific *Oryza sativa* L. x *O. glaberrima* Steud. progenies in upland rice improvement. Euphytica, 92: 237-246.

Linares OF (2002). African rice (*Oryza glaberrima*): History and future potential. PNAS, 99: 16360-16365.

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D, Stein L (2002). Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). DNA Res. 9: 199-207.

Mohapatra S (2010). Pockets of gold. Rice Today, April-June newsletter.

Ndjiondjop MN, Semagn K, Sie M, Cissoko M, Fatondji B, Jones M (2008). Molecular profiling of interspecific lowland rice populations derived from IR64 (*Oryza sativa*) and Tog5681 (*Oryza glaberrima*). Afr. J. Biotechnol. 7(23): 4219-4229.

Nuijten E, van Treuren R, Struik PC, Mokuwa A, Okry F, Teeken B, Richards P (2009) Evidence for the Emergence of New Rice Types of Interspecific Hybrid Origin in West African Farmers' Fields. PLoS ONE 4(10): e7335.

Peeters JP, Martinelli JA (1989). Hierarchical cluster analysis as a tool to manage variation in germplasm collections. Theor. Appl. Genet. 78: 42-48.

Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000). A high-density linkage map of *Theobroma cacao* L. Theor. Appl. Genet. 101: 948-955.

Sarla N, Mallikarjuna Swamy BP (2005). *Oryza glaberrima*: A source for the improvement of *Oryza sativa*. Curr. Sci. 89(6): 955-963.

Second G (1982). Origin of the genetic diversity of cultivated rice (Oryza ssp.): study of the polymorphism scored at 40 isozyme loci. Jpn. J. Genet. 57: 25-57.

Semagn K, Ndjiondjop MN, Lorieux M, Cissoko M, Jones M, McCouch S (2007). Molecular profiling of an interspecific rice population derived from a cross between WAB 56-104 (*Oryza sativa*) and CG 14 (*Oryza glaberrima*). Afr. J. Biotechnol. 6(17): 2014-2022.

Semon M, Nielsen R, Jones MP, McCouch SR (2005). The population structure of African cultivated rice *Oryza glaberrima* (Steud): Evidence for elevated levels of linkage disequilibrium caused by admixture with *O. sativa* and ecological adaptation. Genetics, 169: 1639-1647.

Souza E, Sorrells ME (1991). Relationships among 70 North American oat germplasms: I. Cluster analysis using quantitative characters. Crop Sci. 31: 599-605.

Thiémélé D, Boisnard A, Ndjiondjop MN, Chéron S, Séré Y, Aké S, Ghesquière A, Albar L (2010) Identification of a second major

resistance gene to *Rice yellow mottle virus, RYMV2,* in the African cultivated rice species, *O. glaberrima.* Theor. Appl. Genet. 121(1): 169-179.