Spanish Journal of Agricultural Research
13(2), e0701, 10 pages (2015)
eISSN: 2171-9292
http://dx.doi.org/10.5424/sjar/2015132-6330
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)

RESEARCH ARTICLE

OPEN ACCESS

Genetic relationships between interspecific lines derived from *Oryza glaberrima* and *Oryza sativa* crosses using microsatellites and agro-morphological markers

Yonnelle D. Moukoumbi[1], Olufisayo Kolade[2], Khady N. Drame[3], Moussa Sie[2] and Marie Noëlle Ndjiondjop[2]

[1] *Africa Rice Center (AfricaRice), Sahel Regional Station. BP 96 Saint Louis. Senegal*
[2] *AfricaRice, 01 BP 2031 Cotonou. Benin*
[3] *AfricaRice, Regional Station. Mikocheni B/Kawe. Avocado Street. P.O. Box 33581 Dar-Es-Salaam. Tanzania*

Abstract

New Rice(s) for Africa (NERICA) are high yielding rice varieties mostly cultivated in Sub-Saharan Africa and developed by the Africa Rice Center. This study is aimed at investigating the proportion of introgression of parental genomic contribution of 60 lowland NERICA varieties and establishment of molecular profiling. Agro-morphological data from 17 characteristics was recorded and significant (p<0.05) to high significant (p<0.0001) differences were obtained with leaf length and width, plant height at maturity, days to heading, maturity, primary and secondary branching of panicles, and grain width and grain thickness. A total of 114 microsatellite polymorphic markers covering 2183.13 cM of the rice genome showed the proportions of alleles introgressed from the donor parent (*Oryza glaberrima*) into 52 lowland NERICA lines (TOG5681 and IR64) as follows: 11% for BC_2, 6.07% for BC_3, and 7.55% for BC_4. The introgression proportions for the eight remaining lowland NERICA lines derived from other crosses ranged from 5.5 to 11.3%. The proportion recorded with the recurrent parent was 83.99%. The highest introgression proportions of the *O. glaberrima* allele for all 60 lowland NERICA lines were found on chromosomes 2, 6, and 12 (TOG5681/IR64) and on chromosome 3 with NERIC-L-29 (TOG5681/IR1529-680-3-2). Multivariate analyses performed using an association of agro-morphological and molecular data revealed two major groups according to the distribution of the lowland NERICAs including the lowland NERICAs released were found in cluster 1 of the dendrogram. Genetic and genomic studies, QTL identification and analysis using agro-morphologically significant traits revealed should be used to develop mega-varieties adapted in rice growth conditions in Sub-Saharan Africa.

Additional key words: agro-morphological and SSR markers; introgression; lowland NERICA; rice.

Abbreviations used: DAS (days after sowing); MCA (multiple correspondence analyses); NERICA (New Rice for Africa); NL (Lowland NERICA); PCR (polymerase chain reaction); SSA (Sub-Saharan Africa); SSR (single sequence repeat).

Citation: Moukoumbi, Y. D.; Kolade, O.; Drame, K. N.; Sie, M.; Ndjiondjop, M. N. (2015). Genetic relationships between interspecific lines derived from *Oryza glaberrima* and *O. sativa* crosses using microsatellites and agro-morphological markers. Spanish Journal of Agricultural Research, Volume 13, Issue 2, e0701, 10 pages. http://dx.doi.org/10.5424/sjar/2015132-6330.

Received: 28 May 2014. **Accepted:** 13 Mar 2015.

Copyright © 2015 INIA. This is an open access article distributed under the Creative Commons Attribution License (CC by 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Funding: L'ORÉAL-UNESCO international fellowship for Women in Science.

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to Yonnelle Dea Moukoumbi**:** y.moukoumbi@cgiar.org

Introduction

Rice (*Oryza sativa* L.) is the second most widely grown cereal in the world after wheat. It is one of the main food crops and a staple for the majority of the populations in developing countries. In Sub-Saharan Africa (SSA), the potential of lowland agro-ecosystems to produce rice is much higher than that of upland ecologies, because they are suited to cropping intensification, with the possibility of growing two or more rice crops per year. The development of more lowland ecologies therefore offers a great opportunity for the sustainable expansion of rice production and intensification especially if new rice-based technologies are adopted by farmers. This will contribute to the overall improvement of rice production and increased-incomes for farmers.

The Africa Rice Center (AfricaRice) through the Rice Research and Development Network for West and Central Africa (ROCARIZ) comprised of national agricultural research systems (NARS) of West and Central African countries, has developed high yielding rice varieties that are appropriately adapted to the lowland rice ecologies in Sub-Saharan Africa (SSA), among

them the lowland New Rice for Africa (NERICA). The lowland NERICA lines give real hope for rice productivity improvement, profitability and the sustainability of rice production systems in SSA. They are derived from interspecific crosses between the African rice species (*Oryza glaberrima)* and the Asian rice species (*Oryza sativa indica*). While the African rice species is resistant to diseases and drought, has lower yield potential resulting from high grain shattering and susceptibility to lodging (Jones *et al*., 1997; Futakuchi & Sie, 2009), the Asian rice species has high yield potential. A total of 60 lowland NERICA lines have been developed and include different levels of backcrossing: 4 BC_1, 22 BC_2, 19 BC_3 and 15 BC_4. Several of the released lowland NERICA lines (NL-18, NL-19, NL-20, NL-26, NL-34, NL-39, NL-41, NL-42, NL-49 and NL-60) are widely grown in countries such as Benin (NL-18, NL-19 and NL-20); Burkina Faso (NL-19, NL-20, NL-41 and NL-60), Cameroon (NL-19), Liberia (NL-19), Sierra-Leone (NL-19 and NL-20); Togo (NL-19 and NL-34); Mali (NL-20 and NL-42); and Niger (NL-39 and NL-49). Understanding the genetic variability between the 60 lowland NERICA lines will enable farmers to choose the varieties that are best suited to their ecologies and cultural practices.

By definition, genetic diversity is an inherited variation among and between populations, created, activated and maintained by evolution (Demol *et al.,* 2002). It is a fundamental characteristic without which breeders are very limited and powerless in plant breeding. The study of genetic diversity reposes on adapted and appropriate techniques such as characterization using agro-morphological, physiological, biochemical or molecular markers which have been successfully used in recent years to help in identifying elite promising lines. The 60 lowland NERICA lines have been successfully characterized using agro-morphological markers and showed three clusters, irrespective of the level (4 BC_1, 22 BC_2, 19 BC_3 and 15 BC_4) of the backcross generation (Moukoumbi *et al*., 2011). Eighty percent of the lowland NERICA lines showed characteristics sought in lowland rainfed growth conditions, such as 20-25 tillers per plant (good); an intermediate plant height (110-130cm); early to medium duration (100<days<130); dense secondary panicle branching (approximately 50% of spikelets borne directly on primary branches); and highly fertile panicles (>90%). Molecular characterization has been greatly facilitated by the advent of DNA marker technology in the 1980s, which offered a large number of environmentally-insensitive genetic markers that could be generated to follow the inheritance of important agronomic traits (Peleman & Van der Voort, 2003). Microsatellites (SSR) are co-dominant, multiallelic, highly polymorphic (even in closely related individuals), with high abundance and uniform distribution in plant genomes, and widely used for genetic studies including estimation of the proportions of donor genome and a recurrent parent background (Bernardo *et al*., 2000).

The genotyping of 48 lowland NERICA lines derived from crosses between IR64 (*O. sativa*) and TOG5681 (*O. glaberrima*) using 60 microsatellite markers showed variable proportions of introgression of the *glaberrima* parent (TOG5681) depending on the level of the backcross generation. The estimated averages of donor parent TOG5681 coverage were 7.2% (83.5 cM) at $BC_2F_{10,}$ 8.5% (99.3 cM) at BC_3F_8 and 8.1% (93.8 cM) at BC_4F_8 (Ndjiondjop *et al.,* 2008). As a complement to this first study, the objectives of the present study were (*i*) to estimate the proportion of introgressions from the donor parent; (*ii*) estimate the highest proportion *of O. glaberrima* introgression; (*iii*) establish the molecular profiling of the lowland NERICA lines derived from crosses using TOG5681 and IR64; and (*iv*) assess the genetic relationships among breeding lines quoted above to identify desirable parental combinations, and associate both agro-morphological and molecular SSR markers which could be used for an efficient breeding program.

Material and methods

The planting material used for the study included 60 lowland NERICA (NL-) varieties at different backcross levels. Sixty lowland NERICA lines (NL) derived from TOG5681 (*O. glaberrima*), were used as the donor parent and IR64 (*O. sativa subsp. indica*) as the recurrent parent. The remaining nine NL were derived from other crosses: TOG5681/2*IR64//IR31785-58-1-2-3-3 (NL-53); TOG5681/2*IR64//IR31851-96-2-3-2-1 (NL-23, NL-24 and NL-25); TOG5674/4*IR31785-58-1-2-3-3(NL-43); TOG5675/4*IR28 (NL-47); IR31785//TOG5674/4*IR31785-58-1-2-3-3 (NL-59); TOG5681/4*IR31785-58-1-2-3-3 (NL-22); TOG5681/3*IR1529-680-3-2 (NL-21). The backcross level of the 60 NL- varieties was: 4 BC_1, 22 BC_2, 19 BC_3 and 15 BC_4.

Agro-morphological characterization

The experiments were conducted during the 2008 and 2009 rainy seasons at the AfricaRice experimental station in Ouedeme, located in the southern part of Benin (6°42'46"N, 1°41'07"E, altitude 21 m). An aug-

mented experimental design was laid out in three blocks using $NPK_{15\text{-}15\text{-}15}$ fertilizers as basal application at a rate of 200 kg/ha during land preparation and urea was applied at the rate of 50 kg/ha at 14 days after sowing (DAS) and at panicle initiation. Descriptor data (17 in total) were collected according to descriptors for wild and cultivated rice (*Oryza* spp.) from Bioversity International/IRRI/AfricaRice (2007).

Proportions of introgression

Genomic DNA was extracted from 250 mg of young leaves according to the protocol on mini-preparations (Risterucci *et al*., 2000). Quantification and assessment of DNA quality were performed using a spectrophotometer at 260 nm and 280 nm wavelengths. The genomic DNAs of the 60 NL and all parents were diluted and stored at -20°C.

Genomic DNA was extracted from the 60 lowland NERICA lines and analyzed by simple sequence repeat (SSR) using PCR techniques. The number of primers used ranged from 73 to 250 (Table 1), according to the cross (Orjuela *et al.,* 2010) and 25 µL of total SSR-PCR volume mixture was amplified using the following program: initial denaturation (1 cycle of 94°C for 4 min) followed by 35 amplification cycles including denaturation (94°C for 30 s); hybridization of primers (55°C for 30 s), elongation (72°C for 45 s) and a final elongation (72°C for 5 min). SSR-PCR products were separated on 3% TBE agarose gel electrophoresis with 0.5 X TBE buffer (40 mM Trizma base-HCl, 40 mM boric acid and 1 mM EDTA), stained with 0.5 µg/mL bromophenol blue (3X STR), visualized with ultraviolet Trans-illuminator and the image captured by Alpha Imager HP software. SSR-profiles were scored and analyzed.

For statistical descriptive and variance analyses mixed model was performed from agro-morphological data using XLSTAT (2011) software. Data scoring and statistical analyses were performed as described by Semagn *et al*. (2006). Only clear polymorphic SSR bands of various molecular weight sizes were scored manually in comparison with the respective parents. The letter «A» was attributed to alleles from the donor parent; the letter «B» to alleles from the recurrent parent *(O. sativa*); the letter «H» to heterozygotes; the letter «E» to non-parental alleles; and «-» to signify missing data. The number of polymorphic markers was estimated with Microsoft Excel-2010 for each cross including the 60 lowland NERICAs. The map distances (2183.13 cM) between 114 markers was used as the basis for estimating parental (donor and recurrent) contribution and introgression, the heterozygosity and non-parental genome per chromosome for each lowland NERICA using Graphical GenoTypes (Van Berloo, 2008). The molecular profile was generated with 52 lowland NERICAs derived from crosses between TOG5681 and IR64. Multiple correspondence analyses (MCA) following Ward's (1963) method (XLSTAT, 2011) and cluster analysis using Unweight Neighbors Joining method were carried out to investigate the overall variation and patterns of relationships among lowland NERICA(s) using the

Table 1. Descriptive statistics and probability (α =0.05) of the 17 quantitative traits evaluated in NL-1 to NL-60

Traits	Min	Max	Mean	SD	Probability (α =0.05)
Tillers at 15 DAS (No.)	2.42	8.00	4.61	0.17	0.4465^{ns}
Tillers at 60 DAS (No.)	10.00	28.75	18.78	0.57	0.0674^{ns}
Plant height at maturity (cm)	87.30	178.20	119.57	2.45	0.0129^{*}
Leaf length (mm)	32.38	77.22	47.28	1.13	$<0.0001^{***}$
Leaf width (mm)	1.01	2.53	1.45	0.03	$<0.0001^{***}$
Days to heading	74.00	97.00	86.99	0.61	$<0.0001^{***}$
Maturity (days)	101.00	128.00	110.90	0.79	$<0.0001^{***}$
Panicle primary branching (No.)	6.50	13.90	9.31	0.22	$<0.0001^{***}$
Panicle secondary branching (No.)	12.60	46.30	23.88	0.93	0.0037^{**}
Number of panicles per plant (No.)	5.90	17.40	11.55	0.33	0.6735^{ns}
Panicle length (cm)	21.17	42.10	25.70	0.47	0.0655^{ns}
Spikelet fertility (%)	66.67	96.09	87.40	0.82	0.3568^{ns}
Grain length (mm)	8.91	10.55	9.69	0.05	0.0341^{*}
Grain width (mm)	2.17	2.80	2.43	0.02	0.0341^{*}
Grain thickness (mm)	1.40	2.16	1.93	0.02	$<0.0001^{***}$
1000 grain weight (g)	14.22	33.00	27.13	0.45	0.6094^{ns}
Grain yield (kg/m^2)	2.59	8.64	6.39	0.17	0.1832^{ns}

DAS: days after sowing; SD: standard deviation. *,**,***: significant at 0.05, 0.01-0.001 and 0.0001, respectively; ns: non-significant at 0.05.

software package Graphical GenoTypes (Van Berloo, 1999).

Results

Agro-morphological characterization

Table 1 shows that the differences among agro-morphological traits such as plant height at maturity, panicle secondary branching and grain width were moderately significant ($p<0.05$), while leaf length and width, panicle primary branching, days to heading, maturity and grain thickness were highly significant ($p<0.0001$). Most of the lowland NERICAs were of intermediate plant height (110-130 cm). Indeed, the recorded mean plant height at maturity was 119.57 cm, with a minimum plant height at maturity of 87.30 cm and a maximum of 178.20 cm. Panicle primary and secondary branching was dense, heavy and compact, and rarely open. Leaf width varied between 1.01 and 2.53 mm with a mean of 1.45 mm, while leaf length showed a minimum of 32.38 mm and maximum of 77.22 minimum and maximum. Days to heading observed ranged from early to medium (100<days<130).

Chromosomal repartition of the total number screened and polymorphic markers

The polymorphism survey shown on Table 2 ranged from 23.87% (TOG5681/IR64//IR31851-96-2-3-2-1) to 50.66% (TOG5675/IR28). The number of polymorphic markers per chromosome varied from 2 (chromosome 4) to 17 (chromosome 1).

Allelic contribution of the donor parent (*O. glaberrima*), recurrent parent (*O. sativa*) and non-parental alleles into the 60 lowland NERICA lines

The introgression of the donor parent's (TOG5681) allele into the 60 lowland NERICAs is shown in Table 3. For 52 lowland NERICA varieties, introgression of the donor genome was 11% at BC_2, ranging from 5.6% (NL-2, NL-4 and NL-9) to 15.6% (NL-1, NL-5 and NL-17). At BC_3, introgression ranged from 3.2% (NL-27 and NL-29) to 14.5% (NL-42) with an average of 6.07%. At BC_4, it varied from 5.8% (NL-48, NL-51 and NL-56) to 11.6% (NL-11) with an average of 7.55%. In comparison, the rate of introgression of the recurrent parent's genome ranged from 86.28% (BC_2) to 91.55% (BC_3) and heterozygosity and non-parental alleles were the lowest. For the eight remaining lowland NERICAs, the rate of introgression ranged from 8.2% (NL-22) to 11.3% (NL-21) for the TOG5681 genome; 5.5% (NL-43 and NL-59) for the TOG5674 genome and 9.5% (NL-47) for the TOG5675 genome. The rate of heterozygous introgression was 1.1% in NL-59 (1.1%) and 2.2% in NL-23, NL-24 and NL-25.

High proportion of the introgression of the donor parent *(O. glaberrima)*

The highest proportion of the introgression of the donor parent (*O. glaberrima)* in the 52 NL was on

Table 2. Chromosomal repartition of the total number screened and polymorphic markers per cross

Chr.	TOG 5681/IR64 IR64/TOG5681		TOG5681/ IR1529-680-3-2		TOG5681/IR64// IR31785-58-1-2-3-3		TOG5681/IR64// IR31851-96-2-3-2-1		TOG5674/ IR31785-58-1-2-3-3		TOG5675/ IR28	
	1†	2‡	1	2	1	2	1	2	1	2	1	2
1	28	10	29	10	22	5	25	5	27	15	23	17
2	24	10	28	17	34	11	36	10	28	19	30	12
3	21	11	25	11	28	6	15	3	22	7	20	13
4	16	9	24	9	24	5	17	2	21	12	19	13
5	14	9	20	7	22	6	16	5	17	9	17	12
6	20	9	15	6	17	7	19	4	24	10	18	13
7	17	12	14	7	20	5	15	5	15	3	16	8
8	21	13	19	12	21	9	20	9	21	11	22	11
9	23	7	11	7	16	7	21	5	10	3	16	5
10	14	6	14	8	12	2	20	3	12	5	16	4
11	16	7	14	9	14	5	23	4	8	4	8	3
12	19	11	14	6	20	5	16	3	17	5	24	5
Total	233	114	227	109	250	73	243	58	222	103	229	116

†1: Total number of the screened SSRs; ‡2: Number of the polymorphic SSRs among NERICA-L-1 to NERICA-L-60.

Table 3. Sixty lowland NERICA breeding lines (NL) derived from TOG5681 and IR64 crosses genome coverage

N.º	Lines	Pedigree	Level of backcross	Genome composition (%)				
				TOG5681	IR64	H	NP	MD
1	NL1	TOG5681/3*IR64	BC_2	15.6	83.3	0.0	0.0	1.1
2	NL2	TOG5681/3*IR64	BC_2	5.6	94.4	0.0	0.0	0.0
3	NL3	TOG5681/3*IR64	BC_2	10	88.9	0.0	0.0	0.0
4	NL4	TOG5681/3*IR64	BC_2	5.6	90	0.0	0.0	4.4
5	NL5	TOG5681/3*IR64	BC_2	15.6	77.8	1.1	4.4	1.1
6	NL6	TOG5681/3*IR64	BC_2	8.9	88.9	0.0	1.1	1.1
7	NL7	TOG5681/3*IR64	BC_2	7.8	91.1	0.0	1.1	0.0
8	NL8	TOG5681/3*IR64	BC_2	7.8	90	0.0	1.1	1.1
9	NL9	TOG5681/3*IR64	BC_2	5.6	93.3	0.0	1.1	0.0
10	NL10	TOG5681/3*IR64	BC_2	6.7	88.9	1.1	2.2	0.0
11	NL11	TOG5681/3*IR64	BC_2	14.4	83.3	1.1	0.0	1.1
12	NL12	TOG5681/3*IR64	BC_2	13.3	86.7	0.0	0.0	0.0
13	NL13	TOG5681/3*IR64	BC_2	8.9	90	0.0	1.1	0.0
14	NL14	TOG5681/3*IR64	BC_2	11.1	84.4	1.1	1.1	2.2
15	NL15	TOG5681/3*IR64	BC_2	13.3	82.2	0.0	4.4	0.0
16	NL16	TOG5681/3*IR64	BC_2	13.3	84.4	0.0	2.2	0.0
17	NL17	TOG5681/3*IR64	BC_2	15.6	82.2	0.0	2.2	0.0
18	NL18	TOG5681/3*IR64	BC_2	14.4	80	0.0	4.4	1.1
19	NL19	TOG5681/3*IR64	BC_2	12.2	86.7	0.0	1.1	0.0
20	NL20	TOG5681/3*IR64	BC_2	14.4	82.2	0.0	2.2	1.1
21	NL49	TOG5681/3*IR64	BC_2	11.1	83.3	0.0	2.2	3.3
			Mean BC_2	**11**	**86.28**	**0.21**	**1.51**	**0.83**
22	NL26	TOG5681/4*IR64	BC_3	9.7	83.9	0.0	4.8	1.6
23	NL27	TOG5681/4*IR64	BC_3	3.2	96.8	0.0	0.0	0.0
24	NL28	TOG5681/4*IR64	BC_3	8.1	87.1	1.6	0.0	1.6
25	NL29	TOG5681/4*IR64	BC_3	3.2	96.8	0.0	1.6	0.0
26	NL30	TOG5681/4*IR64	BC_3	4.8	90.3	0.0	0.0	1.6
27	NL31	TOG5681/4*IR64	BC_3	6.5	91.9	0.0	3.2	0.0
28	NL32	TOG5681/4*IR64	BC_3	4.8	95.2	0.0	1.6	0.0
29	NL33	TOG5681/4*IR64	BC_3	6.5	93.5	0.0	0.0	0.0
30	NL34	TOG5681/4*IR64	BC_3	4.8	95.2	0.0	0.0	0.0
31	NL35	TOG5681/4*IR64	BC_3	6.5	93.5	0.0	0.0	0.0
32	NL36	TOG5681/4*IR64	BC_3	9.7	90.3	0.0	0.0	0.0
33	NL37	TOG5681/4*IR64	BC_3	6.5	93.5	0.0	0.0	0.0
34	NL38	TOG5681/4*IR64	BC_3	9.7	87.1	0.0	0.0	1.6
35	NL39	TOG5681/4*IR64	BC_3	4.8	93.5	1.6	1.6	0.0
36	NL40	TOG5681/4*IR64	BC_3	4.8	95.2	0.0	0.0	0.0
37	NL41	TOG5681/4*IR64	BC_3	8.1	88.7	0.0	0.0	1.6
38	NL42	TOG5681/4*IR64	BC_3	14.5	83.9	0.0	0.0	0.0
			Mean BC_3	**6.07**	**91.55**	**0.18**	**0.75**	**0.47**
39	NL44	TOG5681/5*IR64	BC_4	11.6	79.1	0.0	7	2.3
40	NL45	TOG5681/5*IR64	BC_4	8.1	89.5	0.0	1.2	1.2
41	NL46	TOG5681/5*IR64	BC_4	9.3	89.5	1.2	0.0	0.0
42	NL48	IR 64//TOG 5681/4*IR 64	BC_4	5.8	93	0.0	1.2	0.0
43	NL50	IR 64//TOG 5681/4*IR 64	BC_4	7	93	0.0	0.0	0.0
44	NL51	IR 64//TOG 5681/4*IR 64	BC_4	5.8	94.2	0.0	0.0	0.0
45	NL52	IR 64//TOG 5681/4*IR 64	BC_4	8.1	88.4	0.0	1.2	2.3
46	NL53	IR 64//TOG 5681/4*IR 64	BC_4	7	93	0.0	0.0	0.0
47	NL54	IR 64//TOG 5681/4*IR 64	BC_4	7	93	0.0	0.0	0.0
48	NL55	IR 64//TOG 5681/4*IR 64	BC_4	7	93	0.0	0.0	0.0
49	NL56	IR 64//TOG 5681/4*IR 64	BC_4	5.8	93	1.2	0.0	0.0
50	NL57	IR 64//TOG 5681/4*IR 64	BC_4	9.3	86	0.0	3.5	1.2
51	NL58	IR 64//TOG 5681/4*IR 64	BC_4	7	89.5	0.0	1.2	0.0
52	NL60	IR 64//TOG 5681/4*IR 64	BC_4	7	88.4	0.0	2.3	2.3
			Mean BC_4	**7.55**	**90.18**	**0.17**	**1.26**	**0.66**

Table 3 (cont.). Sixty lowland NERICA breeding lines (NL) derived from TOG5681 and IR64 crosses genome coverage.

N.º	Lines	Pedigree	Level of backcross	Genome composition (%)					
				TOG5681	**IR64**	**H**	**NP**	**MD**	
53	NL21	TOG5681/3*IR1529-680-3-2	BC_2	11.3	84.5	0.0	3.1	1.0	
				TOG5681	**IR64**	**IR31785-58-1-2-3-3**	**H**	**NP**	**MD**
54	NL22	TOG 5681 / 2*IR 64 //IR 31785-58-1-2-3-3	BC_1	8.2	3.8	72.2	0.0	6.6	14.8
				TOG5681	**IR64**	**IR31851-96-2-3-2-1**	**H**	**NP**	**MD**
55	NL23	TOG 5681/2*IR 64//IR31851-96-2-3-2-1	BC_1	10.9	2.2	67.4	2.2	17.4	0.0
56	NL24	TOG 5681/2*IR 64//IR31851-96-2-3-2-1	BC_1	8.7	5.0	73.3	2.2	10.9	0.0
57	NL25	TOG 5681/2*IR 64//IR31851-96-2-3-2-1	BC_1	10.9	2.0	74.1	2.2	10.9	0.0
			Mean BC_1	**10.17**	**0.73**	**73.93**	**2.2**	**13.07**	**0.0**
					IR 31785-58-1-2-3-3	**TOG5674**	**H**	**NP**	**MD**
58	NL43	TOG 5674/4*IR 31785-58-1-2-3-3	BC_3		91.2	5.5	0.0	1.1	2.2
59	NL59	IR 31785-58-1-2-3-3// TOG 5674/4*IR31785-58-1-2-3-3	BC_4		91.2	5.5	1.1	2.2	0.0
					TOG5675	**IR28**	**H**	**NP**	**MD**
60	NL47	TOG 5675/4*IR 28	BC_3		9.5	78.4	0.0	12.2	0.0

H: heterozygotes; NP: non parental; MD: missing data.

chromosomes 2, 6 and 12 and the lowest on chromosomes 1 and 4 (Fig. 1). For eight lowland NERICAs (NL-21, NL-22, NL-23, NL-24, NL-25, NL-43, NL-47 and NL-59), the highest introgression occurred on chromosomes 2 (NL-22 and NL-42), 4 (NL-21) and 6 (NL-21 and NL-59). There was no donor parent genome introgression on chromosomes 2, 3, 4, 6, 10 and 12 (NL-23, NL-24 and NL-25).

Lowland NERICA structuring

Cluster and MCA lines are useful to evaluate the potential breeding value of the lowland NERICAs. The first two axes in the MCA explained 90.37% of the total variability (Fig. 2) and revealed the two distinct major groups between lowland NERICA regardless of the level of backcross and the proportion of the introgressed donor parent genome. TOG5681 (*O. glaberrima)* appeared distant from all lowland NERICA lines and other parental lines.

The cluster analysis was performed using the simple matching coefficients derived from 114 SSR markers. The dendrogram produced two distinct clusters with nine sub-clusters (Fig. 3). Forty six percent of the lowland NERICAs were found on cluster 1, including four sub-clusters from NL-1 to NL-34. Eight lowland NERICA varieties released in SSA (NL-18, NL-20, NL-26, NL-34, NL-41, NL-49, NL-54 and NL-60), where a strong variability was found, were included in group 1. Cluster 2 comprised 53.85% of lowland NERICA lines, including NL-19 which is widely cultivated in West African countries and two varieties released in Niger (NL-39) and Mali (NL-42). The major difference observed between the cluster and MCA was the nine sub-groups observed in the cluster analysis, which were not evident in the MCA.

Discussion

Agronomic and morphological traits were examined as recommended by Jacquot & Arnaud (1979) and Glaszmann (1987). The most discriminating quantitative traits were leaf length and width, days to heading, maturity and panicle primary branching. Similar results were reported by Sie (1991) for leaf length and width, grain length and weight as discriminating traits through a study based on genetic evaluation of traditional rice varieties. In addition, these results provide information on the weed suppressive and high yielding characteristics of the 60 lowland NERICAs.

SSR polymorphic markers were well distributed along the 12 rice chromosomes. The main advantages of the SSR markers used are their co-dominance and high polymorphism, even among very closely linked subjects, which showed their efficiency in the assessment of the parental contributions reported by several studies. The estimated *O. glaberrima* genome among interspecific lines (*O. glaberrima* and *O. sativa*) was

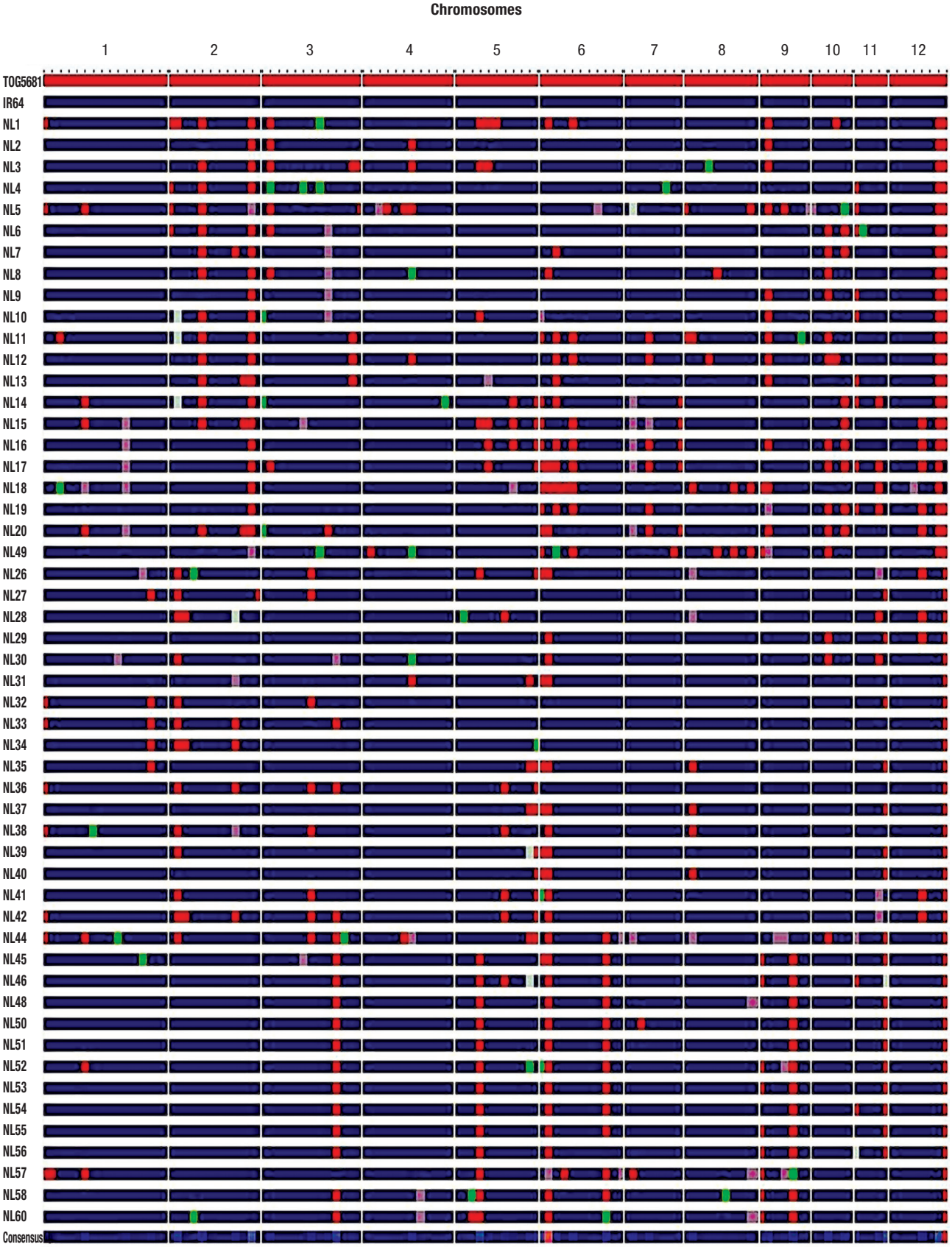

Figure 1. Graphical genotype of 52 lowland NERICA lines (NL) derived from crosses between TOG 5681 and IR64 using 114 microsatellites markers. Vertical bars represent the 12 chromosomes. Each chromosome is represented by horizontal lines with the parent genome (red, *Oryza glaberrima*; dark blue, *Oryza sativa*; light blue, heterozygote; and mauve, non-parental) introgressed. Missing data are shown in green.

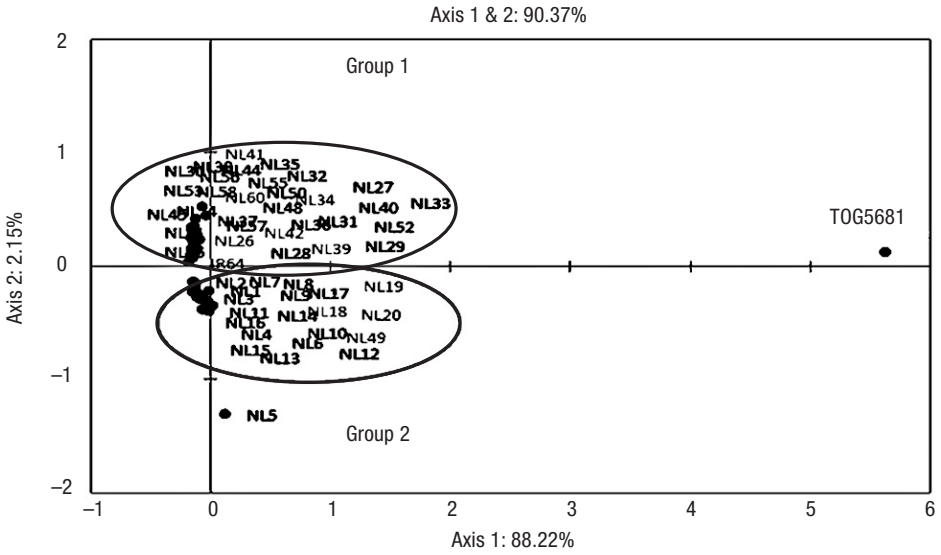

Figure 2. Score plot on the first two principal components from multiple correspondence analyses (MCA) of the 52 lowland lines derived from TOG5681 and IR64 crosses genotyped with 114 SSRs. NERICA lines are described in Table 3.

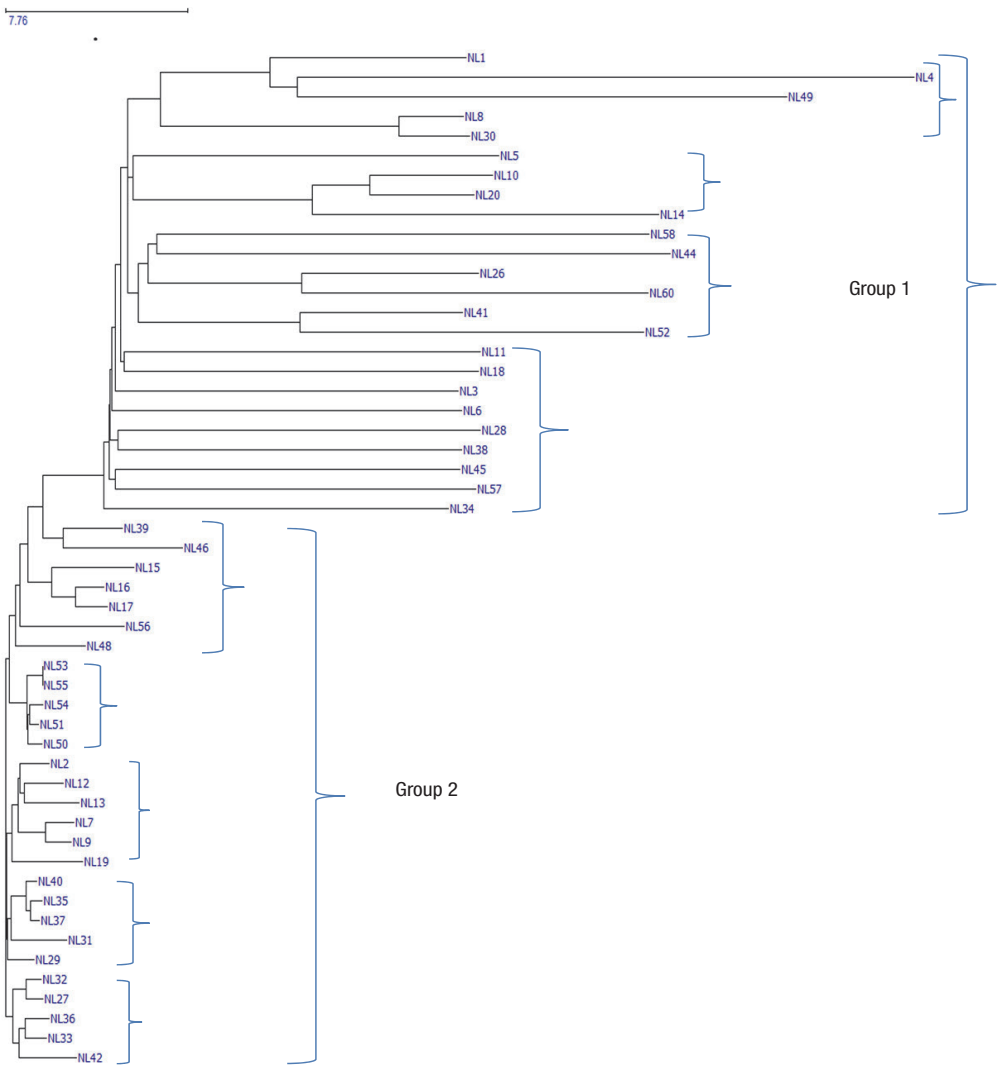

Figure 3. Dendrogram of the 52 NLs using unweight neighbors joining method of clustering derived from 114 microsatellites markers. The 52 NLs were separated into two mean groups and nine sub-groups

7.2, 8.5 and 8.1% at BC_2, BC_3 and BC_4, respectively. In this study, the proportions of the TOG5681 genome were 11, 6.07 and 7.55% at BC_2, BC_3 and BC_4 and statistically different ($p<0.05$). The *O. glaberrima* genome on the other hand was lower at BC_3 than BC_4 and double according to the expected Mendelian inheritance values: 12.5% (BC_2), 6.25% (BC_3) and 3.13% (BC_4). Though Hospital (2005) reported that during successive backcrosses, the genome of the donor parent must move towards zero in all chromosomes except the one carrying the introgressed portion of the allele of interest, the gap observed between expected and estimated donor parent contribution could be explained in different ways. The gap might have resulted from the action of environmental effects on plant growth and development. Phenotypic variability can be observed even between genotypes belonging to the same group with the same parents (Cisse *et al.,* 2006) and between population sizes used for making selections. The number of markers and their distribution could explain the disparities.

Studies using the same parents (Ndjiondjop *et al*., 2008; Agnoun *et al.,* 2012) reported the highest proportion of the introgression on chromosome 6. The study showed that the highest introgression occurred on chromosomes 2 and 12. Variability of the introgressed donor parent genome has been widely observed. The lack of introgression of the donor parent genome in some lowland NERICA lines could explain some phenotypical differences observed during vegetative and reproductive stages of these lowland NERICAs. Indeed, the genome of the *O. glaberrima* parent can be partially introgressed on progenies when the crosses are carried out with an *O. sativa* variety (Barry *et al.,* 2007). Also intensive selection occurred during the selfing of the BC_2, BC_3 and BC_4 generations according to the number of the traits concerned. As mentioned by Heckenberger *et al.* (2005), selection and genetic drift during inbreeding might explain the differences observed between the current and expected proportions of the donor parent genome in the 60 lowland NERICA varieties. The proportion of introgression of the donor parent genome is not the critical factor but rather what it represents in terms of genetic information, including the number of the genes accumulated. The lack of donor parent genomes on chromosomes 2, 3, 4, 5, 6, 10, and 12 in NL-23, NL-24 and NL-25 derived from the crossing carried out with three parents might be justified because these varieties have never gone beyond experimental selection in spite of their demonstrated tolerance to salinity and cold. Seed admixtures, spontaneous mutation, anther sterility during pollination resulting from sporo-gametophytic (Zeng *et al.,* 2009) during the development of the 60 lowland NERICA varieties could justify the presence of the non-parental alleles. Regarding the lowland NERICA varieties NL-43, NL-47 and NL-59 derived from crosses with TOG5674 and TOG5675, the proportion of introgression was highest and ranged from 5.5 to 9.5%. These three lowland NERICAs varieties might have introgressed the resistance genes *rymv1-4 and rymv1-5* alleles (Albar *et al.,* 2003; 2006) which were identified in TOG5674 and TOG5675 (*O. glaberrima* varieties).

The 10 lowland NERICA varieties released in Sub-Saharan Africa are found in the nine sub- clusters and eight of them belong to cluster 1, where 80% of lowland NERICA varieties showed characteristics adapted to lowland rice growing conditions. Cluster 2 included lowland NERICA varieties that might be grown in both upland and lowland ecologies. Lowland NERICA-5 derived from TOG5681/IR64 might be screened for some abiotic and biotic stresses and might reveal other quantitative trait loci (QTLs) for tolerance or resistance genes hidden in the TOG5681 variety. This information can help other researchers to identify important agronomic traits and encourage research on QTLs for various stresses. Molecular analysis shows a wider genome of the 60 lowland NERICA lines than the agro-morphological analysis reported by previous studies.

References

Agnoun Y, Sie M, Djedatin G, Drame KN, Toulou B, Ogunbayo SA, Sanni KA, Tia D, Ahanchede A, Vodouhe RS, Ndjiondjop MN, 2012. Molecular profiling of interspecific lowland rice progenies resulting from crosses between TOG5681 and TOG5674 *(Oryza glaberrima)* and IR64 *(Oryza sativa)*. Int J Biol 4: 3-7. http://dx.doi.org/10.5539/ijb.v4n3p19.

Albar L, Ndjiondjop MN, Esshak Z, Berger A, Pinel A, Jones M, Fargette D, Ghesquiere A, 2003. Fine genetic mapping of a gene required for Rice yellow mottle virus cell-to-cell movement. Theor Appl Genet 107: 371-378. http://dx.doi.org/10.1007/s00122-003-1258-4.

Albar L, Bangratz-Reyser M, Hebrard E, Ndjiondjop MN, Jones M, Ghesquiere A, 2006. Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to Rice yellow mottle virus. Plant J 47: 417-426. http://dx.doi.org/10.1111/j.1365-313X.2006.02792.x.

Barry MB, Pham JL, Noyer JL, Billot C, Courtois B, Ahmadi N, 2007. Genetic diversity of the two cultivated rice species *(O. sativa* and *O. glaberrima)* in Maritime Guinea: Evidence for interspecific recombination. Euphytica 154: 127-137. http://dx.doi.org/10.1007/s10681-006-9278-1.

Bernardo R, Romero-Severson J, Ziegle J, Hauser J, Joe L, Hookstra G, Doerge RW, 2000. Parental contribution and coefficient of coancestry among maize inbreds: pedigree, RFLP, and SSR data. Theor Appl Genet 100: 552-556.

Bioversity International/IRRI/AfricaRice, 2007. Descriptors for wild and cultivated rice *(Oryza* spp.), pp: 21-43. Available in http://www.bioversityinternational.org/uploads/tx_news/Rice_232_01.pdf.

Cisse F, Yalcouye N, Doumbia Y, Hamadoun A, 2006. Development of interspecific populations *O. sativa* L. and *O. glaberrima* Steud and identification of strains resistant to drought conditions in lowland rice cultivation and controlled flooding. Proc Beyond the first generation NERICA in Africa: paradigms and partnerships for the next decade. AfricaRice, Dar es Salaam. pp: 23-24.

Demol J, Baudoin JP, Louant BP, Maréchal R, Mergeai G, Otoul E, 2002. Amélioration des plantes. Application aux principales espèces cultivées en régions tropicales [Plant breeding: application to the main species grown in tropical regions]. Presses Agronomiques, Gembloux, Belgique. 581 pp.

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

Glaszmann JC, 1987. Isozymes and classification of Asian rice varieties. Theor Appl Genet 74: 21-30. http://dx.doi.org/10.1007/BF00290078.

Heckenberger M, Bohn M, Frisch M, Maurer HP, Melchinger AE, 2005. Identification of essentially derived varieties with molecular markers: an approach based on statistical test theory and computer simulations. Theor Appl Genet 111: 598-560. http://dx.doi.org/10.1007/s00122-005-2052-2.

Hospital F, 2005. Selection in backcross programmes. Philos T R Soc B 360: 1503-1511. http://dx.doi.org/10.1098/rstb.2005.1670.

Jacquot M, Arnaud M, 1979. Classification numérique de variétés de riz. Agro Trop 34: 157-173.

Jones MP, Dingkuhn M, Aluko GK, Semon M, 1997. Interspecific *Oryza sativa* L. × *O. glaberrima* Steud. progenies in upland rice improvement. Euphytica 92: 237-246. http://dx.doi.org/10.1023/A:1002969932224.

Moukoumbi YD, Sie M, Vodouhe R, Ogunbayo SA, N'dri B, Toulou B, Ahanchede A, 2011. Assessing phenotypic diversity of interspecific rice varieties using agro-morphological characterization. J Plant Breed Crop Sci 3: 74-86.

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: 4219-4229.

Orjuela J, Garavito A, Bouniol M, Arbelaez JD, Moreno L, 2010. A universal core genetic map for rice. Theor Appl Genet 120: 563-572. http://dx.doi.org/10.1007/s00122-009-1176-1.

Peleman JD, Van der Voort JR, 2003. Breeding by design. Trends Plant Sci 8: 330-334. http://dx.doi.org/10.1016/S1360-1385(03)00134-1.

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. http://dx.doi.org/10.1007/s001220051566.

Semagn K, Ndjiondjop MN, Cissoko M, 2006. Microsatellites and agronomic traits for assessing genetic relationships among 18 New Rice for Africa (NERICA) varieties. Afr J Biotechnol 5: 800-810.

Sie M, 1991. Prospection and genetic evaluation of rice traditional varieties *(Oryza sativa* L et *O. glaberrima* Steud) from Burkina Faso. PhD Thesis. Faculty of Sciences and Technics of National University of Ivoiry Coast, Abidjan. 118 pp.

Van Berloo R, 2008. GGT 2.0: Versatile software for visualization and analysis of genetic data. J Hered 99: 232-236. http://dx.doi.org/10.1093/jhered/esm109.

Ward JH, 1963. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58: 236-244. http://dx.doi.org/10.1080/01621459.1963.10500845.

XLSTAT, 2011. XLSTAT Statistical Software for Excel. Available in www.xlstat.com.en.

Zeng YX, Hu CY, Lu YG, Li JQ, Liu XD, 2009. Abnormalities occurring during female gametophyte development result in the diversity of abnormal embryo sacs and leads to abnormal fertilization in *indica/japonica* hybrids in rice. J Integr Plant Biol 51 (1): 3-12. http://dx.doi.org/10.1111/j.1744-7909.2008.00733.x.