Full Length Research Paper

Molecular profiling of interspecific lowland rice populations derived from IR64 (Oryza sativa) and Tog5681 (Oryza glaberrima)

Marie Noelle Ndjiondjop¹*, Kassa Semagn², Moussa Sie¹, Mamadou Cissoko¹, Blandine Fatondji¹ and Monty Jones³

¹Africa Rice Center (WARDA), 01 BP 2031, Cotonou, Benin.
²International Maize and Wheat Improvement Center (CIMMYT), P. O. Box 1041, Nairobi, Kenya.
³Forum for Agricultural Research in Africa, PMB CT 173 Cantonments, Accra, Ghana.

Accepted 24 October, 2008

Several lowland NERICAs (New Rice for Africa) were derived from crosses between IR64 (an Oryza sativa subsp. indica variety) and Tog5681 (an Oryza glaberrima variety) that possess useful traits adapted to lowland conditions in West Africa. The proportion of parental genomic contribution and extent of genetic differences among these sister lines is unknown at the molecular level. The objectives in this study were therefore to determine, with 60 SSR markers that cover 1162 cM of the rice genome, the frequency and magnitude of deviations from the expected parental contributions among 21 BC²F¹⁰, 17 BC³F⁸ and 10 BC⁴F⁸ lines and determine patterns of their genetic relationships. The estimated average O. glaberrima genome coverage was 7.2% (83.5 cM) at BC²F¹⁰, 8.5% (99.3 cM) at BC³F⁸ and 8.1% (93.8 cM) at BC⁴F⁸ lines. The O. sativa parent accounted for 73.2% (851.3 cM) at BC²F¹⁰, 82.6% (959.5 cM) at BC³F⁸ and 78.2% (908.6 cM) at BC⁴F⁸. Non-parental alleles were detected at all 3 backcross generations but the frequency of such alleles at BC²F¹⁰ (8.8%) was twice that of BC³F⁸ (3.4%) and nine times that of BC⁴F⁸ (0.9%). Both cluster and principal component analyses revealed two major groups irrespective of the level of backcross generations and the proportion of parental genome contribution.

Key words: African rice, introgression, microsatellite, NERICA, Oryza glaberrima, rice.

INTRODUCTION

In areas of high population density, African farmers are intensifying cropping on fragile upland soils leading to severe environmental degradation and loss in production potential. The development of appropriate lowland rice technologies is one of the strategies to help farmers to reduce pressure on degraded uplands by shifting cultivation to the lowlands. To improve African rice production, breeders at the Africa Rice Center (WARDA) and the National Agricultural Research Systems (NARS) have continued in developing interspecific rice varieties by crossing the two cultivated species, Asia rice (Oryza sativa) and Africa rice (Oryza glaberrima). Various studies have shown that very useful agronomic traits exist within the traditional landraces of O. glaberrima. These include (i) rapid and profuse vegetative growth coupled with droopy lower leaves (Jennings et al., 1979) that contribute to weed competitiveness (Fofana et al., 1995); (ii) moderate to high levels of resistance to blast (Silue and Notteghem, 1991), rice yellow mottle virus (Attere and Fatokun, 1983; John et al., 1985), rice gall midge (Sauphanor, 1985; Alam, 1988) and nematodes (Reversat and Destombes, 1995); (iii) reasonably good levels of tolerance to soil acidity, iron toxicity and drought (Sano et al., 1984; Jones et al., 1994). The grain of O. glaberrima also has good aroma and taste. However, cultivation of O. glaberrima is being increasingly replaced by O. sativa varieties due to low yield-performance, high shattering and susceptibility to lodging (Jones et al., 1997). Interspecific hybridization is an important tool to transfer desirable genes between species (Roy, 1984; Katiyar et al., 1998), generate morpho-physiological variation.

*Corresponding author. E-mail: m.ndjiondjop@cgiar.org.
(Choudhary and Joshi, 2001), elucidate intergenomic relationships (Mizushima, 1980; Attia et al., 1987; Choudhary et al., 2000), and develop synthetic amphiploids (Olsson, 1960). Non-homologous pairing between different genomes (Mizushima, 1980; Choudhary et al., 2000) provides immense opportunities for generating variability and broadening the gene pool through interspecific hybridization.

Earlier efforts to transfer useful genes between *O. sativa* and *O. glaberrima* have been frustrated by high spikelet sterility in the progeny. In the early 1990s, WARDA developed over 3000 interspecific BC$_2$ inbred lines from a cross between an upland *O. sativa* tropical japonica variety, WAB 56-104, as the recipient parent, and an *O. glaberrima* variety, CG14, as the donor parent. Several upland NERICA (New Rice for Africa) varieties were released for production in different parts of Africa. The most popular upland NERICA varieties combine the best traits of both parents: high yield presumably derived from the *O. sativa* parent and the ability to thrive in harsh environments from the *O. glaberrima* parent (Jones et al., 1997; www.warda.org). The NERICA varieties offer real hope for improving the productivity, profitability, and sustainability of rice farming in sub-Sahara Africa. WARDA and NARS breeders then began making crosses between *O. sativa indica* (IR64) and *O. glaberrima* (Tog5681) to create interspecific lines adapted to lowland conditions. This work has resulted in the generation of many stable and high yielding interspecific lowland NERICA varieties that are well adapted to lowland rice production systems of the region. The proportion of parental genomic contribution and extent of genetic differences among these lowland sister lines is unknown at the molecular level.

Three recent advances have rapidly changed the face of rice genomic research: the use of microsatellite or simple sequence repeats (SSRs), the availability of functional markers such as expressed sequence tags (ESTs), and the completion of the rice genomic sequence (International Rice Genome Sequencing Project, 2005). Microsatellites are considered ideal markers for genetic studies because they are co-dominant, multiallelic, highly polymorphic even in closely related individuals, high abundance and uniform distribution in plant genomes (Morgante and Olivieri, 1993; Rafalski and Tingey, 1993; Sharma et al., 1995; Brondani et al., 1998). The sequence information and map positions of rice SSR markers are publicly available (http://www.gramene.org) and more rice SSR markers are being developed to tag any possible polymorphic parents (Temnykh et al., 2000; McCouch et al., 2002; International Rice Genome Sequencing Project, 2005). Many research groups have used SSR markers for various purposes, including estimation of the proportion of donor genome in a recurrent parent background (Bernardo et al., 2000; Heckenberger et al., 2005b; Ndjidjop et al., 2006; Semagn et al., 2007). Semagn et al. (2007) used SSR markers to examine the proportion of introgression and extent of genetic differences among 70 BC$_2$ interspecific upland rice lines developed by crossing a *tropical japonica* variety (WAB 56-104) as the recurrent parent to an *O. glaberrima* variety (CG 14) as the donor parent. That study revealed an average introgression of 6.3% from the *O. glaberrima* parent and a recovery of 87.4% of the *O. sativa* recurrent parent. The objectives of the present study were to (a) estimate the extent of introgressions from the donor, and the degree and rate of recovery of the recurrent parent alleles among 21 BC$_2$F$_{10}$, 17 BC$_3$F$_8$ and 10 BC$_4$F$_8$ interspecific lines developed by crossing *O. glaberrima* (Tog5681) as donor parent and *O. sativa* subsp. *indica* (IR64) as recurrent parent, and (b) determine the genetic relationships among these 48 lines and evaluate the potential breeding value of the lines that have not yet been released as variety.

**MATERIAL AND METHODS**

**Plant materials**

Sixty lowland NERICA lines, developed by crossing an *O. sativa* subsp. *indica* variety IR64 as recurrent parent and *O. glaberrima* variety Tog5681 as donor parent, were selected by WARDA Varietal Nomination Committee. Forty eight of these interspecific lines (Table 1) were used in this study. IR64 is an irrigated improved Asian variety with several desirable agronomic traits, such as high yield, short growth duration and plant height adapted to irrigated conditions. Tog5681 is a variety from Nigeria and it has low yield potential due primarily to grain shattering and susceptibility to lodging (Jones et al., 1997). However, Tog5681 has several useful traits, including long panicle length, high weed-competitiveness ability as a result of early vigor and high tiller number, resistance to rice yellow mottle virus (Ndjidjop et al., 1999), and resistance to nematodes *Heterodera sacchari* (Lorieux et al., 2000). All lines in this study had gone either through two generations (21 lines), three generations (17 lines) or four generations (10 lines) of backcrossing followed by eight to ten generations of selfing for fixing the lines (Table 1).

**Genotyping and data analyses**

DNA was extracted from 2 weeks old seedlings using a Cetyl trimethylammonium bromide (CTAB) protocol as described by Saghai-Maroof et al. (1984). One hundred and forty SSR primers were selected according to their position on the genetic map by Temnykh et al. (2001) and screened for polymorphism between parents. Sixty polymorphic markers (Table 2) were then used for genotyping the 48 interspecific lines. Polymerase chain reactions (PCR), PCR product separation, data scoring and statistical analyses were performed as described by Semagn et al. (2007). Four different statistical analyses were performed. First, the proportion of the donor and recurrent parental genome was calculated using the software package GGT, 2007 edition (http://www.dpw.wau.nl/pub/ggt; Van Berloo, 1999). Map distances in centiMorgans (abbreviated as cM) between markers were used as the basis for estimating parental contribution per chromosome and line. Second, statistical differences in the proportion of parental genome contributions, heterozygosity and non-parental alleles among the 3 backcross generations were examined using one-way analysis of variance (ANOVA) and tested with LSD for pairwise comparisons using SPSS version 12 (SPSS, Inc.). Third, simple matching coefficients (the ratio of number of matches to total number markers) were calculated as a measure of
Table 1. Pedigree and parental contribution for 21 BC$_2$F$_1$, 17 BC$_3$F$_8$ and 10 BC$_4$F$_8$ interspecific lines developed using an *O. sativa* variety IR64 and an *O. glaberrima* variety Tog5681. Lines released as varieties are indicated in boldface, with variety name given in parenthesis in the first column. Expected donor/recurrent parent contributions with Mendelian inheritance for BC$_2$, BC$_3$ and BC$_4$-derived inbred lines are 12.5/87.5, 6.25/93.75% and 3.13/96.88%, respectively.

<table>
<thead>
<tr>
<th>Line name (variety name)</th>
<th>WARDA designation</th>
<th>Pedigree</th>
<th>Level of backcross</th>
<th>Genome composition (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tog 5681</td>
</tr>
<tr>
<td>WAS 122-IDSA 10-WAS 1-1 FKR 1</td>
<td>NERICA-L-1</td>
<td>TOG5681/IR64/2* IR64</td>
<td>BC$_2$</td>
<td>11.7</td>
</tr>
<tr>
<td>WAS 122-IDSA 10-WAS 6-1 FKR 1</td>
<td>NERICA-L-2</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>6.6</td>
</tr>
<tr>
<td>WAS 122-IDSA 11-WAS 11-4-FKR 1</td>
<td>NERICA-L-3</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>5.7</td>
</tr>
<tr>
<td>WAS 122-IDSA 11-WAS 8-2</td>
<td>NERICA-L-4</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>7.7</td>
</tr>
<tr>
<td>WAS 122-IDSA 12-WAS B-FKR 1</td>
<td>NERICA-L-5</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>6.2</td>
</tr>
<tr>
<td>WAS 122-IDSA 13-WAS 10-FKR 1</td>
<td>NERICA-L-6</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>10.8</td>
</tr>
<tr>
<td>WAS 122-IDSA 13-WAS 13-3-3 FKR 1</td>
<td>NERICA-L-7</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>16.2</td>
</tr>
<tr>
<td>WAS 122-IDSA 14-WAS B- FKR 1</td>
<td>NERICA-L-8</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>7.9</td>
</tr>
<tr>
<td>WAS 122-IDSA 10-WAS-3-1-TGR 3</td>
<td>NERICA-L-9</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>7.1</td>
</tr>
<tr>
<td>WAS 122-IDSA 10-WAS-7-2-FKR1-TGR 8</td>
<td>NERICA-L-10</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>7.7</td>
</tr>
<tr>
<td>WAS 122-IDSA 11-WAS -10-2-TGR 60</td>
<td>NERICA-L-11</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>7.4</td>
</tr>
<tr>
<td>WAS 122-IDSA 11-WAS -B-IER-11-19</td>
<td>NERICA-L-12</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>1.6</td>
</tr>
<tr>
<td>WAS 122-IDSA -13-WAS 10- WAB-B-TGR 5</td>
<td>NERICA-L-13</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>6.6</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS 2-WAB 1-TGR 6</td>
<td>NERICA-L-14</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>5.0</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS -2</td>
<td>NERICA-L-15</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>3.4</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS -2-B-1-TGR 132</td>
<td>NERICA-L-16</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>7.6</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS -2-WAB2-TGR 7</td>
<td>NERICA-L-17</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>1.1</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS -4-B-1-TGR 121</td>
<td>NERICA-L-18</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>10.9</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS -6-1 (FKR62N)</td>
<td>NERICA-L-19</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>8.2</td>
</tr>
<tr>
<td>WAS 122-IDSA -1-WAS -B (FKR60N and N2)</td>
<td>NERICA-L-20</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>8.0</td>
</tr>
<tr>
<td>WAS 122-IDSA-1-B-IER-18-6 (Niger)</td>
<td>NERICA-L-49</td>
<td>TOG5681/IR64/2* IR64</td>
<td>&quot;</td>
<td>3.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean for BC$_2$</th>
<th>7.2</th>
<th>73.2</th>
<th>0.9</th>
<th>8.8</th>
<th>9.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAS 161-B-1-1-FKR 1</td>
<td>NERICA-L-26</td>
<td>TOG5681/IR64/3* IR64</td>
<td>BC$_3$</td>
<td>8.8</td>
<td>82.0</td>
</tr>
<tr>
<td>WAS 161-B-2-B-1</td>
<td>NERICA-L-27</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>10.3</td>
<td>81.1</td>
</tr>
<tr>
<td>WAS 161-B-2-B-2</td>
<td>NERICA-L-28</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>8.1</td>
<td>82.3</td>
</tr>
<tr>
<td>WAS 161-B2-B3</td>
<td>NERICA-L-29</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>8.6</td>
<td>80.3</td>
</tr>
<tr>
<td>WAS 161-B-2-B-4</td>
<td>NERICA-L-30</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>7.6</td>
<td>85.9</td>
</tr>
<tr>
<td>WAS 161-B-4-1-FKR 1</td>
<td>NERICA-L-31</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>11.7</td>
<td>81.3</td>
</tr>
<tr>
<td>WAS 161-B-4-B-1-TGR 51</td>
<td>NERICA-L-32</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>8.1</td>
<td>87.6</td>
</tr>
<tr>
<td>WAS 161-B-4-B-2</td>
<td>NERICA-L-33</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>10.3</td>
<td>87.8</td>
</tr>
<tr>
<td>WAS 161-B--6-3-FKR 1</td>
<td>NERICA-L-34</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot;</td>
<td>6.0</td>
<td>90.4</td>
</tr>
</tbody>
</table>
genetic similarity between the 48 lines and used to generate dendrogram using the complete-link method of SAHN clustering. Fourth, principal component analysis (PCA) was used to investigate the overall variation and patterns of relationship among the lines. Both cluster and principal component analyses were performed using NTSYS-pc for windows, version 2.0, Exeter Software (Rohlf, 1998).

## RESULTS

### Marker polymorphism and parent contribution

An initial polymorphism survey conducted using DNA from the two parents (IR64 and Tog5681) revealed a polymorphism level of 77.9% (109 out of the 140 SSR markers). The number of polymorphic markers per chromosome varied from 7 on chromosome 10 to 11 on chromosomes 7, and the overall average was 9.1 polymorphic markers per chromosome (data not shown). Sixty (Table 2) out of the 109 polymorphic markers were selected for genotyping the 48 interspecific lines and the average number of markers used for genotyping was 5 per chromosome.

Knowledge on the proportion of donor parent contribution among lines and chromosomes provides useful information for selection and variety development. In our study, *O. glaberrima* DNA varied from 1.1 to 16.2% of the genome at BC2F10, 3.3 to 12.3% at BC3F8 and 1.1 to 13.3% at BC4F8 (Table 1). The average proportion of the genome containing *O. glaberrima* alleles at BC2F10, BC3F8 and BC4F8 was 7.2% (83.5 of 1,162.1 cM), 8.5% (99.3 of 1,162.1 cM) and 8.1% (93.8 cM), respectively. The introgression of *O. glaberrima* DNA across chromosomes was highly variable. There was no *O. glaberrima* DNA on (i) chromosomes 3, 5 and 11 at BC2F10, (ii) chromosomes 3, 4, 5, 9, 11 and 12 at BC3F8, and (iii) chromosomes 4, 5, 9, 11 and 12 at BC4F8 (Figure 1). Chromosome 6 contained the highest *O. glaberrima* DNA at all the 3 backcross generations (54.8% at BC2F10, 70.9% at BC3F8 and 88.1% at BC4F8). *O. sativa* genome varied from 52.6 to 90.1% at BC2F10, from 62.0 to 90.4% at BC3F8 and from 53.4 to 90.7% at BC4F8. The average proportion of the *O. sativa* parent was 73.3% (851.3 of 1,162.1 cM) at BC2F10, 82.6% (959.5 cM) at BC3F8 and 78.2% (905.0 cM) at BC4F8.

<table>
<thead>
<tr>
<th>WAS 161-B-6-4-FKR 1</th>
<th>NERICA-L-35</th>
<th>TOG5681/IR64/3* IR64</th>
<th>&quot; 11.1</th>
<th>82.2</th>
<th>0.0</th>
<th>0.7</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAS 161-B-6-B1</td>
<td>NERICA-L-36</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 6.5</td>
<td>82.1</td>
<td>0.0</td>
<td>4.5</td>
<td>6.9</td>
</tr>
<tr>
<td>WAS 161-B-6-B-4</td>
<td>NERICA-L-37</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 11.1</td>
<td>84.1</td>
<td>0.0</td>
<td>2.2</td>
<td>2.6</td>
</tr>
<tr>
<td>WAS 161-B-6-B-1-B</td>
<td>NERICA-L-38</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 8.5</td>
<td>85.1</td>
<td>0.0</td>
<td>4.1</td>
<td>2.2</td>
</tr>
<tr>
<td>WAS 161-B-6-WAB-B-TGR 16 (Niger)</td>
<td>NERICA-L-39</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 12.3</td>
<td>80.3</td>
<td>1.2</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>WAS 161-B-9-1-FKR 1</td>
<td>NERICA-L-40</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 3.3</td>
<td>85.3</td>
<td>0.0</td>
<td>6.1</td>
<td>5.4</td>
</tr>
<tr>
<td>WAS 161-B-9-3 (FKR58N)</td>
<td>NERICA-L-41</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 4.4</td>
<td>62.0</td>
<td>7.6</td>
<td>16.4</td>
<td>9.6</td>
</tr>
<tr>
<td>WAS 161-IDSA-3-WAS-B-IER-2-4 (N1)</td>
<td>NERICA-L-42</td>
<td>TOG5681/IR64/3* IR64</td>
<td>&quot; 8.6</td>
<td>83.8</td>
<td>0.8</td>
<td>2.2</td>
<td>4.6</td>
</tr>
</tbody>
</table>

| Mean for BC3 | 8.5 | 82.6 | 0.6 | 3.4 | 4.9 |

<table>
<thead>
<tr>
<th>WAS 191-10-3-FKR 1</th>
<th>NERICA-L-48</th>
<th>IR64/TOG5681/4*IR64</th>
<th>BC4 10.6</th>
<th>79.0</th>
<th>0.0</th>
<th>1.6</th>
<th>8.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAS 191-10-4-FKR 1-TGR 123</td>
<td>NERICA-L-50</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 10.1</td>
<td>74.2</td>
<td>0.0</td>
<td>0.0</td>
<td>15.6</td>
</tr>
<tr>
<td>WAS 191-10-WAB-B-TGR 23</td>
<td>NERICA-L-51</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 1.1</td>
<td>64.5</td>
<td>0.0</td>
<td>0.0</td>
<td>34.4</td>
</tr>
<tr>
<td>WAS 191-1-5-FKR 1</td>
<td>NERICA-L-52</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 7.6</td>
<td>84.1</td>
<td>0.0</td>
<td>0.0</td>
<td>8.3</td>
</tr>
<tr>
<td>WAS 191-4-10</td>
<td>NERICA-L-54</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 11.4</td>
<td>82.9</td>
<td>0.0</td>
<td>0.0</td>
<td>5.7</td>
</tr>
<tr>
<td>WAS 191-8-1-FKR 1</td>
<td>NERICA-L-55</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 7.6</td>
<td>90.7</td>
<td>0.0</td>
<td>0.0</td>
<td>1.7</td>
</tr>
<tr>
<td>WAS 191-8-3</td>
<td>NERICA-L-56</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 7.6</td>
<td>90.7</td>
<td>0.0</td>
<td>0.0</td>
<td>1.7</td>
</tr>
<tr>
<td>WAS 191-9-B-2</td>
<td>NERICA-L-57</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 13.3</td>
<td>53.4</td>
<td>4.1</td>
<td>5.3</td>
<td>23.8</td>
</tr>
<tr>
<td>WAS 191-9-WAB-B-TGR 24</td>
<td>NERICA-L-58</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 6.5</td>
<td>71.7</td>
<td>3.3</td>
<td>0.0</td>
<td>18.5</td>
</tr>
<tr>
<td>WAS 191-9-3- FKR-1 (FKR58N)</td>
<td>NERICA-L-60</td>
<td>IR64/TOG5681/4*IR64</td>
<td>&quot; 4.9</td>
<td>90.5</td>
<td>0.0</td>
<td>1.9</td>
<td>2.7</td>
</tr>
</tbody>
</table>

| Mean for BC4 | 8.1 | 78.2 | 0.7 | 0.9 | 12.1 |
Table 2. Genetic map distance (Temnykh et al., 2001) of the 60 SSR markers used in the present study.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Marker name</th>
<th>Map distance (cM)</th>
<th>Chromosome</th>
<th>Marker name</th>
<th>Map distance (cM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RM84</td>
<td>26.4</td>
<td>7</td>
<td>RM180</td>
<td>27.5</td>
</tr>
<tr>
<td>1</td>
<td>RM220</td>
<td>28.4</td>
<td>7</td>
<td>RM501</td>
<td>34.7</td>
</tr>
<tr>
<td>1</td>
<td>RM259</td>
<td>54.2</td>
<td>7</td>
<td>RM11</td>
<td>47.0</td>
</tr>
<tr>
<td>1</td>
<td>RM486</td>
<td>153.5</td>
<td>7</td>
<td>RM429</td>
<td>96.9</td>
</tr>
<tr>
<td>1</td>
<td>RM315</td>
<td>165.3</td>
<td>8</td>
<td>RM152</td>
<td>9.4</td>
</tr>
<tr>
<td>2</td>
<td>RM555</td>
<td>34.7</td>
<td>8</td>
<td>RM25</td>
<td>52.2</td>
</tr>
<tr>
<td>2</td>
<td>RM424</td>
<td>67.5</td>
<td>8</td>
<td>RM72</td>
<td>60.9</td>
</tr>
<tr>
<td>2</td>
<td>RM475</td>
<td>92.5</td>
<td>8</td>
<td>RM404</td>
<td>69.0</td>
</tr>
<tr>
<td>2</td>
<td>RM573</td>
<td>143.7</td>
<td>8</td>
<td>RM223</td>
<td>80.5</td>
</tr>
<tr>
<td>2</td>
<td>RM425</td>
<td>168.1</td>
<td>8</td>
<td>RM230</td>
<td>112.2</td>
</tr>
<tr>
<td>2</td>
<td>RM208</td>
<td>188.4</td>
<td>8</td>
<td>RM433</td>
<td>116.0</td>
</tr>
<tr>
<td>3</td>
<td>RM60</td>
<td>0.0</td>
<td>9</td>
<td>RM409</td>
<td>45.6</td>
</tr>
<tr>
<td>3</td>
<td>RM007</td>
<td>64.0</td>
<td>9</td>
<td>RM257</td>
<td>66.1</td>
</tr>
<tr>
<td>3</td>
<td>RM554</td>
<td>100.6</td>
<td>9</td>
<td>RM553</td>
<td>76.7</td>
</tr>
<tr>
<td>3</td>
<td>RM16</td>
<td>131.5</td>
<td>9</td>
<td>RM160</td>
<td>82.4</td>
</tr>
<tr>
<td>4</td>
<td>RM471</td>
<td>53.8</td>
<td>9</td>
<td>RM215</td>
<td>99.4</td>
</tr>
<tr>
<td>4</td>
<td>RM564</td>
<td>73.1</td>
<td>10</td>
<td>RM474</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>RM119</td>
<td>76.1</td>
<td>10</td>
<td>RM239</td>
<td>25.2</td>
</tr>
<tr>
<td>4</td>
<td>RM470</td>
<td>115.5</td>
<td>10</td>
<td>RM467</td>
<td>46.8</td>
</tr>
<tr>
<td>4</td>
<td>RM348</td>
<td>135.4</td>
<td>10</td>
<td>RM294</td>
<td>87.1</td>
</tr>
<tr>
<td>5</td>
<td>RM13</td>
<td>31.4</td>
<td>11</td>
<td>RM167</td>
<td>37.5</td>
</tr>
<tr>
<td>5</td>
<td>RM164</td>
<td>80.1</td>
<td>11</td>
<td>RM479</td>
<td>50.6</td>
</tr>
<tr>
<td>5</td>
<td>RM26</td>
<td>122.7</td>
<td>11</td>
<td>RM536</td>
<td>55.1</td>
</tr>
<tr>
<td>5</td>
<td>RM480</td>
<td>130.6</td>
<td>11</td>
<td>RM209</td>
<td>73.9</td>
</tr>
<tr>
<td>6</td>
<td>RM508</td>
<td>0.0</td>
<td>11</td>
<td>RM229</td>
<td>77.8</td>
</tr>
<tr>
<td>6</td>
<td>RM589</td>
<td>3.2</td>
<td>11</td>
<td>RM244</td>
<td>120.1</td>
</tr>
<tr>
<td>6</td>
<td>RM204</td>
<td>25.1</td>
<td>12</td>
<td>RM20</td>
<td>3.2</td>
</tr>
<tr>
<td>6</td>
<td>RM253</td>
<td>37.0</td>
<td>12</td>
<td>RM558</td>
<td>10.0</td>
</tr>
<tr>
<td>7</td>
<td>RM481</td>
<td>0.8</td>
<td>12</td>
<td>RM247</td>
<td>32.3</td>
</tr>
<tr>
<td>7</td>
<td>RM125</td>
<td>24.8</td>
<td>12</td>
<td>RM17</td>
<td>109.1</td>
</tr>
</tbody>
</table>

and 78.2% (908.6 cM) at BC₄F₈ (Table 1). The highest proportion of *O. sativa* alleles was observed on chromosome 9 both at BC₂F₁₀ (95.2%) and BC₃F₈ (94.2%) and chromosome 5 at BC₃F₈ (97.5%). The highest frequency of heterozygosity per line was 4.2% at BC₂F₁₀, 7.6% at BC₃F₈ and 4.1% at BC₄F₈. Heterozygosity at BC₂F₁₀, BC₃F₈ and BC₄F₈ accounted for 0.9% (10.2 of 1,162.1 cM), 0.6% (7.2 cM of 1162.1 cM) and 0.7% (8.6 of 1,162.1 cM) of the total genome, respectively. A total of 5 chromosomes at BC₂F₁₀, 3 chromosomes at BC₃F₈ and 2 chromosomes at BC₄F₈ contained heterozygous loci. The frequencies of non-parental alleles among lines derived from BC₂F₁₀, BC₃F₈ and BC₄F₈ were 0-28.4, 0-16.4 and 0-5.3%, respectively. The average frequency of overall non-parental alleles per line was 8.8% (101.8 of 1,162.1 cM) at BC₂F₁₀, 3.4% (39.4 of 1,162.1 cM) at BC₃F₈ and 0.9% (10.2 of 1,162.1 cM) at BC₄F₈ (Table 1). Non-parental alleles at least in one line were observed in a total number of 9, 11 and 5 chromosomes at BC₂F₁₀, BC₃F₈ and BC₄, respectively (Figure 1). The highest proportion of non-parental alleles was observed on chromosome 5 (23.2%) at BC₂F₁₀ and chromosome 6 both at BC₃F₈ (28.0%) and BC₄F₈ (10.9%).

ANOVA revealed significant differences among the 3 backcross generations only for the recurrent parent genome (p < 0.04) and non-parental alleles (p < 0.02). Pairwise comparison of means using LSD indicated the presence of significantly higher (p = 0.011) recurrent parent genome at BC₃F₈ than BC₂F₁₀. In contrast, the proportion of non-parental alleles at BC₂F₁₀ was significantly higher than those of BC₃F₈ (p = 0.027) and BC₄F₈ (p = 0.07).

**Genetic similarity and relationship**

Similarity indices and patterns of relationships among
lines from cluster and Principal component analyses are useful to evaluate the potential breeding value of the lines that have not yet been released as variety. The lowest genetic similarly (55.6%) was obtained between NERICA-L-11 and NERICA-L-28 and NERICA-L-28 and NERICA-L-51, NERICA-L-28 and NERICA-L-56, NERICA-L-31 and NERICA-L-37, NERICA-L-51 and NERICA-L-52, NERICA-L-1 and NERICA-L-55, NERICA-L-51 and NERICA-L-55, and NERICA-L-56 and NERICA-L-56 (data not shown). These similarities were also evident in the cluster analysis performed using the simple matching coefficients derived from SSR markers. The phenogram produced two major groups and six sub-groups (Figure 2). All except the first sub-group were represented by at least 1 released variety. The first five principal components (PCs) from principal component analysis explained 54.4% of the variations. A plot of PC1 (22.9%) and PC2 (9.5%) from the principal component analysis revealed the two major groups (Figure 3) in same way as the cluster analysis. There were, however, two differences between the cluster and PCA: (i) the six sub-groups observed in the cluster analysis were not evident in the PCA; and (ii) NERICA-L-17 in group 2 and NERICA-L-3, NERICA-L-39 and NERICA-L-41 in group-1 appeared to be distant from all others within the same group.

DISCUSSION

Hundreds of microsatellite markers have been incorporated into rice genetic maps constructed using both intraspecific and interspecific populations (Chen et al., 1997; Cho et al., 2000; Lorieux et al., 2000; Temnykh et al., 2001). For many studies, large numbers of markers are not required. Rather, a well-distributed, highly informative and robust set of markers would be of particular value. Most of the microsatellite markers used in the present study were well-distributed along the rice chromosomes. However, there were a few intervals with uneven distribution of markers, primarily due to the lack of polymorphic markers within those intervals.

Figure 1. Pie charts for 12 rice chromosomes depicting the proportion of genome introgression among (a) 21 BC$_2$F$_{10}$, (b) 17 BC$_3$F$_8$, and (c) 10 BC$_2$F$_8$ interspecific lowland population derived from Tog5681 (donor) and IR64 (recurrent) parents. The pie charts were plotted from the graphical genotyping analyses outputs. Numbers in the center of the pies correspond to the number of SSR markers used in the study.
Figure 2. Dendrogram of the 48 interspecific lines using simple matching coefficient derived from 60 microsatellite markers. Each line is named with a prefix number to indicate WARDA designation as shown in Table 1 followed by level of backcross (BC$_2$, BC$_3$, or BC$_4$). The 7 released varieties are indicated either with their variety name or country of release as suffix; FKR56N, FKR58N, FKR60N, and FKR62N all are released in Burkina Faso; two varieties with N1 and N2 as suffix are released in Mali.

Molecular markers provide a useful means for estimating parental contributions to inbred progeny (Lorenzen et al., 1995; Visscher, 1996; Bernardo et al., 1997, 2000; Heckenberger et al., 2005a; Frisch and Melchinger, 2006; Semagn et al., 2007). Lorenzen et al. (1995), for example, used restriction fragment length polymorphism (RFLP) markers to detect significant differences between estimated and expected parental contributions in 4 out of 24 soybean (Glycine max) varieties. Bernardo et al. (1997) used RFLP markers to determine the frequency and magnitude of deviations from the expected parental contribution among F$_2$- and BC$_1$-derived maize (Zea
mays) inbreds. Among 34 BC-derived maize inbreds, eight had estimates that deviated significantly from the expected contribution of 75% from the recurrent parent. Bernardo et al. (2000) reported the superiority of SSR markers over RFLP markers in estimating parental contribution. In the present study, the average proportions of genome containing the *O. glaberrima* parent at BC₂, BC₃ and BC₄ were 7.2, 8.5 and 8.1%, respectively, and were not statistically different. However, the *O. glaberrima* genome at BC₂ was significantly (p < 0.05) lower than what would be expected (12.5%) for a random set of BC₂ lines and the expected (6.25% at BC₃ and 3.13% at BC₄).

In another study, Semagn et al. (2007) used microsatellite markers for estimating the contribution of each parent among 70 BC₂ interspecific lines developed for upland climatic conditions. The average recurrent parent (*O. sativa* variety WAB56-104) genome in the upland interspecific lines was consistent with what was expected for a random set of BC₂ lines (87.5%) but the average proportion of introgressed *O. glaberrima* genome (6.3%) was half of what would be expected at BC₂ generation (12.5%). Therefore, the estimated donor parent genome at BC₂ in the present study is in agreement with Semagn et al. (2007). However, *O. glaberrima* genome both at BC₃ and BC₄ as well as the *O. sativa* genome at all 3 backcross generations disagrees with Semagn et al. (2007). About 82% of the BC₃ lines and 90% of the BC₄ lines (Table 1) showed an introgression higher than expected for random lines at the same generation. On the other hand, three BC₂F₁₀ lines and none of the BC₃F₈ and BC₄F₈ lines contained *O. sativa* alleles greater than the expected value at the same generation.
two possible explanations for such deviations between the estimated and expected parental contribution. First, intensive selection during selfing at both BC$_3$ and BC$_4$ generations might have been done in favor of the *O. glaberrima* parent for a number of traits which could accumulate several genes from the donor parent. Second, selection and genetic drift during inbreeding might have caused differences between the actual and expected proportions of the genome derived by an inbred from each of its parents (St. Martin, 1982; Lorenzen et al., 1995; Visscher, 1996; Bernardo et al., 1997, 2000; Heckenberger et al., 2005b; Frisch and Melchinger, 2006). The extent of recovery of recurrent parent genome may depend on the difference between the agronomic performances of the starting donor and recipient parents. The donor *O. glaberrima* parent was of poor agronomic performance and only used to introduce a particular trait to the *O. sativa* parent. In wide crosses and/or with undesirable linkages, therefore, a greater number of backcrosses may be necessary to increase the recovery of the recurrent parent genome other than the target regions. Usually 6 backcrosses with selection for type in the early generations have proved sufficient although Young and Tanksley (1989) found one cultivar developed after 11 backcrosses still containing the entire chromosome arm carrying the gene from the donor parent, which is the case on chromosome 6 in this study.

The average genome containing non-parental alleles per line was the lowest at BC$_3$F$_8$ (0.9%) and the highest at BC$_2$F$_{10}$ (8.8) while those of the BC$_3$F$_6$ were intermediate (3.4%; Table 1). Non-parental alleles have been reported in other studies (Smith et al., 1997; Bernardo et al., 2000; Semagn et al., 2007; Liu et al., 2008). Smith et al. (1997) gave several possible reasons for the existence of non-parental bands (alleles). First, residual heterozygosity may be present in the parental inbreds. A problem with marker analysis of inbreds and their parents is that DNA samples are taken from plants grown from remnant seed stocks. Ideally, DNA samples should be taken from the actual plants used to make the original cross from which an inbred was developed. Residual heterozygosity may therefore cause marker genotypes to differ between remnant seed stocks and the actual plants used to make the cross. Second, contamination by stray pollen may have occurred during inbred development. Third, seed stocks of inbreds may have changed genetically over time though spontaneous mutations at SSR loci or physical mixing with seed from another inbred parent or undocumented outcrossing during generation advance. However, the reduction in non-parental alleles with increasing number of backcross generations in this study suggests that contamination by stray pollen at the early generation of selfing as the main cause for non-parental alleles in this study. The highest frequencies of non-parental alleles were observed on chromosome 5 at BC$_2$ and chromosome 6 both at BC$_3$ and BC$_4$ (Figure 1). The reason for the high frequency of non-parental alleles on chromosome 5 is unclear but chromosome 6 is known to contain a sporo-gametophytic sterility gene in the vicinity of the waxy starch synthase gene (Sano, 1990; Lorieux et al., 2000; Heuer and Miezan, 2003). Sterility association with this locus could affect the production of viable pollen within a line and would tend to promote outcrossing.

The 3 released lowland NERICA (FKR60N, FKR62N and NERICA-L-49) were selected among lines that belong to group-2 (Figures 2 and 3). All the other five subgroups in the cluster analysis that consisted of a total of 41 lines were represented only by 4 released varieties, suggesting that the other lines may provide an opportunity for selection and additional varietal development. However, the possibility for further selection and varietal development within group-1 will be highly dependent on the availability of reliable morpho-agronomic data from both multi-location trials and participatory varietal selections. This study is the first attempt to characterize the introgression of chromosome segments among interspecific lines developed for the lowland climatic conditions in Africa and provides valuable information for breeders.

**ACKNOWLEDGEMENT**

This work was supported by USAID project.

**REFERENCES**


