Grain Protein Content of Interspecific Progenies Derived from the Cross of African Rice (*Oryza glaberrima* Steud.) and Asian Rice (*Oryza sativa* L.)

Hideo Watanabe*, Koichi Futakuchi, Monty P. Jones** and Benjamin A. Sobambo

(Africa Rice Center (WARDA), 01B.P. 2031 Cotonou, Benin)

Abstract: The protein contents of the grain of 50 interspecific progenies developed from the cross between WAB56-104, an *Oryza sativa* variety, and CG 14, an *Oryza glaberrima* line, were investigated. In contrast to the higher protein content of *O. glaberrima* than *O. sativa* on the average, the protein content of CG 14 was always lower than that of WAB56-104. However, judging from the average of three seasons, 72% of the interspecific progenies had a higher protein content than the mid-parent and 50% of them had a higher protein content than WAB56-104. Although the actual values of protein content of the interspecific progenies were significantly different among the seasons, a highly significant correlation was always observed in protein content between any two of the three seasons. Protein content therefore was considered character of each interspecific progeny though it was also affected by environment. A significant correlation was not observed between paddy yield and protein content in any season; several interspecific progenies showed higher protein content and paddy yield than the mid-parents. A low paddy yield is likely to be associated with high protein content through physiological regulation without a genetic linkage between the two traits. However, the results suggest that the transgressive segregation of protein content observed in the interspecific progenies is attributed not to this physiological regulation but to a certain mechanism to concentrate protein in grains with a genetic background.

Key words: African rice, Interspecific hybridization, *Oryza glaberrima* Steud., *Oryza sativa* L., Protein content, West Africa.

African rice (*Oryza glaberrima* Steud.) is the cultivated *Oryza* species different from Asian rice (*Oryza sativa* L.). *O. glaberrima* was domesticated in West Africa more than 3500 years ago (Jones et al., 1997b) and is still cultivated in this sub-region (Singh et al., 1997). Since the yield level of *O. glaberrima* is generally low compared to that of *O. sativa* due to grain shattering and poor resistance to lodging, its growing area has been decreasing. However, many *O. glaberrima* lines are known to have resistance to various biotic and abiotic stresses in West Africa (Jones et al., 1997b; Johnson et al., 1998; Jones and Singh, 1999; Futakuchi et al., 2001) and be important as genetic resources to develop rice varieties suitable for West African resource-poor farmers, who are suffering from low yielding due to multiple stresses in their fields. Juliano and Villareal (1993) have reported that protein content of grain in several *O. glaberrima* lines is higher than that of many *O. sativa* varieties. We also identified *O. glaberrima* lines with a high protein content by screening more than 150 accessions in the Africa Rice Center (WARDA) in comparison with several *O. sativa* varieties in West Africa (Watanabe et al., 2004). *O. glaberrima* is also noteworthy in this aspect because in West African countries the supply of protein is still low though that of food is almost sufficient on a calorie basis (FAO, 1999).

To combine characteristics of *O. sativa* favorable to yield formation and resistance of *O. glaberrima* to the major constraints in West Africa, WARDA commenced interspecific crossing between *O. sativa* and *O. glaberrima* in 1992 and succeeded in developing fixed fertile interspecific progenies (Jones et al., 1997a). In WARDA, some agronomic traits in several interspecific lines of the first generation (WAB450 series) were examined (Jones et al., 1996; Jones et al., 1997a; Dingkuhn et al., 1997b; Dingkuhn et al., 1998; Dingkuhn et al., 1999a; Dingkuhn et al., 1999b). WAB450 series is the progeny of the cross between WAB56-104 (an *O. sativa* upland improved variety) and CG 14 (*O. glaberrima*). The productivity of 22 interspecific progenies (WAB450 series) was examined using the parents and other *O. sativa* varieties as check varieties, all of which are popular varieties in Côte d’Ivoire, under high and low input levels of fertilizer management (Jones et al., 1997a). The high input management comprised basal N-P2O5-K2O application at a rate of 20-36-36 kg ha⁻¹ and two times of top-dressed N at a rate of 40 kg ha⁻¹.
significantly higher yield (5.6 t ha\(^{-1}\)) with necessary hand weeding to keep the fields clean. In the low input management, which simulated farmers’ practices, only nitrogen was applied at a rate of 20 kg ha\(^{-1}\) at 23 and 45 days after seeding with two times of hand weeding at 21 and 42 days after seeding. In the high input conditions, 11 interspecific progenies were the top yielder with average yields of 3.4 to 3.8 t ha\(^{-1}\). They significantly outyielded the \(O.\) glaberrima parent, CG14, and were equivalent to the best \(O.\) sativa check, IDSA 6, in yield. In low input conditions, CG 14 had significantly higher yield than IDSA 6, and 2 interspecific progenies also outyielded IDSA 6. In another experiment, one interspecific progeny out of the 4 WAB450 series tested showed a significantly higher yield (5.6 t ha\(^{-1}\)) than the \(O.\) sativa parent, WAB56-104, with the 120 kg ha\(^{-1}\) of nitrogen application, whereas there was no significant yield difference between the 4 interspecific progenies and WAB56-104 under the other nitrogen levels, i.e. 0, 40, and 80 kg ha\(^{-1}\) (Jones et al., 1997a). The high protein content of \(O.\) glaberrima will be valuable for improving the nutrition state in nutrient-poor regions, and an in-depth evaluation of the interspecific progenies for protein content should be valuable as a protein source.

In \(O.\) sativa, the effect of environments such as cultivating conditions on protein content is large (Hillerislambers et al., 1973; Higashi et al., 1974). However, a tangible varietal difference is observed (Webb et al., 1968), and the rank of varieties in relation to protein content does not differ so much between different cropping seasons and cultivating conditions though the absolute values of protein content highly vary (Matsue et al., 1996; Matsue and Ogata, 2000). In grains of \(O.\) sativa varieties with a small sink size, namely low yield, the accumulation of starch is more disturbed than that of protein because amino acids are firstly translocated to grains and sucrose later (Higashi et al., 1974), suggesting that low yield is likely to be associated with high grain protein content without any genetic linkage between the two traits. Because of this physiological regulation, a negative correlation seems to be observed sometimes between yield and protein content (Kataoka, 1978). However, the varieties having a high protein content without reduced yield were identified, and they were considered to have some mechanism to concentrate protein in the grains with a genetic background (Kataoka, 1973; Higashi et al., 1974, 1976; Tsuzuki and Furusho, 1986). Such information on \(O.\) sativa is not yet available for the interspecific progenies. In the present study, the protein content of the interspecific progenies was compared with that of their parents and a popular practical \(O.\) sativa variety to seek interspecifics progenies with a high protein content. To evaluate the compatibility of high yield and high protein content in interspecific rice from \(O.\) sativa \(\times\) \(O.\) glaberrima crosses, we also examined the correlation of protein content with paddy yield.

Materials and Methods

1. Plant materials

Interspecific \(O.\) sativa \(\times\) \(O.\) glaberrima progenies newly developed by WARDA were used. Before WARDA started the interspecific breeding of \(O.\) sativa \(\times\) \(O.\) glaberrima, 1130 accessions of \(O.\) glaberrima were screened for early maturity, high tillering ability and rapid seedling growth in 1991 and 1992, and eight lines were nominated for wide hybridization. These eight \(O.\) glaberrima lines selected were expected to compete with weeds and escape drought and late-season fungal disease. Grain quality traits other than grain shape such as protein content were not evaluated during this selection. As \(O.\) sativa parents for the wide crosses, five high yielding elite japonica type upland lines developed by WARDA, all of which were officially released in Côte d’Ivoire and/or other West African countries later on, were nominated. Crossing was made in 1992. Seeds of only seven crosses that produced a few fertile grains were collected, and F\(_2\) progenies were successively backcrossed to the respective \(O.\) sativa parents. After two backcrossings, individuals from the BC\(_2\)F\(_1\) were subjected to pedigree selection in upland with selection criteria such as high tillering, rapid seedling growth, early maturity, resistance to lodging and panicle types. Protein content was not included in the criteria. At the BC\(_2\)F\(_1\) generation, seed fertility was between 30 and 65% and progenies of only two crosses, WAB449 and WAB450, showed increased fertility in each generation. The selection continued until varietal traits were fixed after additional six generations. Thus promising interspecific progenies were selected from WAB450 series. Details of the selection of the interspecific progenies are mentioned in Jones et al. (1997a, 1997b).

Fifty lines of the interspecific progenies (WAB450 series) were used for the trial. All of them were derived from the cross between WAB56-104, an \(O.\) sativa variety, and CG 14, an \(O.\) glaberrima line. WAB56-104, which was also developed at WARDA, is a japonica type improved upland variety and officially released in Côte d’Ivoire in 1998 (MINAGRA, 1998). CG 14 was collected from a farmers’ upland field in Casamanse, Senegal, and it was adaptable to lowland conditions (Futakuchi et al., 2001) as well as to upland conditions. Like all other \(O.\) glaberrima lines, CG 14 has not been modified genetically by breeders. We also used Bouake 189, an indica-type improved lowland variety, as a check in addition to the parents because this is one of the most popular varieties in Côte d’Ivoire.

The experiment was conducted and repeated in three consecutive years in the irrigated lowland rice field at WARDA’s research station at M’bé in Côte d’Ivoire (7º52’N, 5º6’W and altitude 300 m). Seedlings were raised on a semi-irrigated nursery bed.
with the standard nursing practice of WARDA in the wet seasons of 1997 (1997 WS) and 1998 (1998 WS) and in the dry season of 1999 (1999 DS). The seeding date was 14 July, 28 July and 10 February in 1997 WS, 1998 WS and 1999 DS, respectively. The seedlings were transplanted 21 days after seeding at a rate of two seedlings per hill with a spacing of 0.25 m × 0.25 m in clean-weeded plots, based on an incomplete block design with three replications. The plots measured 5 m × 5 m (19 hills × 19 rows) and were fertilized with a basal application of compound fertilizer (10% N, 18% P2O5, and 18% K2O) at a rate of 10 g m⁻². Urea (46%N) was top-dressed at a rate of 8.70 g m⁻² at the maximum tillering and panicle initiation stages. The plots were irrigated and manually weeded, as needed to keep the plots clean. To test the potential performance of the interspecific progenies in relation to protein content and yield, we applied the standard varietal evaluation interspecific progenies in relation to protein content.

### Table 1. Summary of the analysis of variance of the effect of lines and cropping seasons on grain protein content in the interspecific progenies.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sum of square</th>
<th>Degree of freedom</th>
<th>Variance</th>
<th>Ratio of variance</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line</td>
<td>38.71</td>
<td>46</td>
<td>0.84</td>
<td>4.10</td>
<td>***</td>
</tr>
<tr>
<td>Season</td>
<td>61.02</td>
<td>2</td>
<td>30.60</td>
<td>149.23</td>
<td>***</td>
</tr>
<tr>
<td>Error</td>
<td>18.87</td>
<td>92</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>118.78</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***: significant at the 0.1% probability level.

2. Determination of grain protein content

Paddy samples collected from the three plots were mixed together and winnowed. The paddy was hulled by Satake Testing Husker (THU 35H) and milled with a milling machine for laboratory use (Yamamoto Test Rice Whitener VP-31T). The milled rice was ground and powdered with a Udy Cyclone Mill with a 1.0 mm mesh screen (WARDA Grain Quality Laboratory, 1999). Total nitrogen content was determined with a near infrared reflectance (NIR) analyzer using the powdered samples. Protein content was calculated by multiplying total nitrogen content by 5.95, the constant to convert nitrogen content to protein content in rice, and expressed on a dry matter basis of milled rice. Since protein is more concentrated in the surface layer of endosperm than in the core of endosperm (Kambayashi et al., 1984), milling yield (a percentage ratio of milled rice to brown rice on a weight basis) may have a positive correlation with protein content of milled rice. In the present experiments, however, no positive correlation was observed between them in any season, indicating that milling yield did not affect the results. For milling yield of the interspecific progenies, the average and standard errors were 86.9% and 1.6%, respectively, in 1997 WS; 87.6% and 1.4%, respectively, in 1998 WS; and 84.6% and 2.0%, respectively, in 1999 DS.

Three lines were eliminated from the analysis because grains of some different line were contaminated in the paddy samples of these three lines in one of the three experimental years. Therefore, 47 lines were tested for protein content.

### Results

1. Protein content of interspecific progenies

In the analysis of variance, both line and cropping season had highly significant effects on protein content of grains (Table 1). According to the variance ratios, the effect of cropping season, namely environment, was large but the difference among lines was also clearly observed in protein content.

In the frequency distribution of protein content (Fig. 1), some of the interspecific progenies showed higher values than WAB56-104 (OSP), the higher protein content parent, in the three seasons. The average of a percentage ratio of the interspecific progenies...
showing higher protein content than WAB56-104 was 50% through the three seasons (Table 2). The 72% of the interspecific progenies had higher protein content than the mid-parent on the average for the three seasons (Table 2). The interspecific progenies showed transgressive segregation in this trait.

A large number of interspecific progenies had a higher protein content than the popular O. sativa variety, Bouake 189, i.e., 75, 47 and 100% in 1997 WS, 1998 WS and 1999 DS, respectively (Fig. 1).

A coefficient of correlation between protein content and 1000-grain weight was -0.139, -0.253 and -0.360 in 1997 WS, 1998 WS and 1999 DS, respectively. In the last two years, the correlation was significant, of which probability level was P < 0.05 in 1998 WS and P < 0.01 in 1999 DS.

2. Comparison of the protein content of the interspecific progenies between different cropping seasons

Although the protein content of the grains of interspecific progenies was significantly different between the seasons (Table 1), a significant correlation (P < 0.001) was observed between any two of the three seasons (Table 3). A similar result was obtained in Spearman’s rank correlation though the probability level of significance was P < 0.01 between 1997 WS and 1999 DS (Table 3). These data indicate that the rank of the interspecific progenies in protein content is considerably maintained between different cropping seasons even though the absolute values of protein content greatly vary with the season. This suggests that protein content is a character of respective interspecific progenies.

3. Correlation of paddy yield with the protein content

No significant correlation was observed between paddy yield and protein content in the interspecific progenies in any three seasons (Fig. 2). In 1997 WS and 1998 WS, most interspecific progenies showed higher yield and protein content than the mid-parents, though only a few interspecific did so in 1999 DS (Fig. 2). The transgressive segregation in protein content observed in the interspecific progenies (Fig. 1 and Table 2) was not attributed to the physiological regulation due to a low sink capacity, namely low paddy yield. These results suggest that interspecific progenies with a high protein content have some mechanism to highly concentrate protein in their grains.

Discussion

The protein content of CG 14 was always lower than that of WAB56-104 through the three seasons (Fig. 1 and Table 2). We expected high protein content of this O. glaberrima line prior to the experiment because of the report by Juliano and Villareal (1993) showing high protein content of O. glaberrima compared to O. sativa. In another experiment, the range of protein content of O. glaberrima ranged from 8.30 to 14.9% (Watanabe et al., 2004), whereas that of CG 14 was 8.33% on the average of the three seasons in the present experiment. Although experimental conditions were not exactly same in the two experiments, CG 14 is an exceptional O. glaberrima line in this aspect. However, transgressive segregation was always observed in the interspecific progenies during the three seasons (Fig. 1 and Table 2). One example of those interspecific progenies showing higher
protein content than the parents was WAB450-B-1A1.1 whose protein content was 9.21, 9.59 and 11.79%, in 1997 WS, 1998 WS and 1999 DS, respectively. A highly significant correlation was always observed in protein content of the interspecific progenies between different cropping seasons (Table 3), and high protein content of some interspecific progenies was not associated with low paddy yield (Fig. 2), which sometimes makes protein content high through physiological regulation (Higashi et al., 1974). Thus we suppose that the transgressive segregation in protein content in the interspecific progenies is not a seeming phenomenon but has a genetic background.

Although the interspecific progenies were developed as upland lines, the materials were raised in irrigated lowland. This is to compare protein content of the interspecific progenies between different climate conditions, wet and dry seasons, without drought damage and check if protein content is a stable character of each interspecific progeny. However, zinc deficiency suppressed the growth of the interspecific progenies in 1998 WS. Therefore, high yielding interspecific progenies were not found in 1998 WS (Fig. 2). According to a rice catalogue for varieties released in Côte d’Ivoire (MINAGRA, 1998), the yield level of WAB56-104, the O. sativa parent, is 3.0 t ha\(^{-1}\) and potential yield of this variety is 4.0 t ha\(^{-1}\) (MINAGRA, 1998). WAB56-104 is not a low yielding variety among them. In our experiment, yield of WAB56-104, always fell in an expected range that is from 3.0 to 4.0 t ha\(^{-1}\) (Fig. 2). On the other hand, many of the interspecific progenies had higher yield than WAB56-104, and some lines with the yield of above 5 t ha\(^{-1}\) were found in 1997 WS and 1999 DS (Fig. 2). The yield level of the interspecific progenies is high compared to existing upland varieties released in Côte d’Ivoire. CG 14 had similar or high yield compared with WAB56-104 (Fig. 2). Low yield of O. glaberrima is due to grain shattering and its yield potential estimated by the product of spikelet number is not inferior to O. sativa (Dingkuhn et al., 1998). CG 14 sometimes showed higher yield than O. sativa varieties inclusive of WAB56-104 in lower fertilizer input conditions where less lodging, consequently less grain shattering, occurs (Jones et al., 1997a; Dingkuhn et al., 1998). In higher fertilizer input conditions, CG 14 showed better dry matter growth and had a larger number of spikelets than WAB56-104, but lower yield probably due to grain shattering (Dingkuhn et al., 1998). In the present trial, we attempted to reduce the number of shattered grains by using cotton sacks, and possibly owing to reduced grain shattering, CG 14 had a higher yield than WAB56-104’s.

To examine the possibility of developing higher

---

**Table 2. Relationship between the interspecific progenies and their parents in the frequency distribution of grain protein content.**

<table>
<thead>
<tr>
<th>Season</th>
<th>Below the lower parent (%)</th>
<th>Between the parents (%)</th>
<th>Above the higher parent (%)</th>
<th>Above the average of the parents (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>97WS</td>
<td>19 (CG 14)</td>
<td>15</td>
<td>66 (WAB56-104)</td>
<td>75</td>
</tr>
<tr>
<td>98WS</td>
<td>9 (CG 14)</td>
<td>40</td>
<td>51 (WAB56-104)</td>
<td>83</td>
</tr>
<tr>
<td>99DS</td>
<td>6 (CG 14)</td>
<td>62</td>
<td>32 (WAB56-104)</td>
<td>60</td>
</tr>
<tr>
<td>mean</td>
<td>11</td>
<td>39</td>
<td>50</td>
<td>72</td>
</tr>
</tbody>
</table>

A variety or line name in a parenthesis indicates the corresponding parent in each season. The WS and DS indicate wet and dry seasons, respectively.

---

**Table 3. Coefficients of correlation of grain protein content between different cropping seasons.**

<table>
<thead>
<tr>
<th>Season</th>
<th>Regression equation</th>
<th>Correlation coefficient</th>
<th>Spearman’s rank correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>97WS vs 98WS</td>
<td>Y = 0.59X + 4.11</td>
<td>0.512 **</td>
<td>0.512 ***</td>
</tr>
<tr>
<td>97WS vs 99DS</td>
<td>Y = 0.76X + 3.62</td>
<td>0.423 **</td>
<td>0.401 *</td>
</tr>
<tr>
<td>98WS vs 99DS</td>
<td>Y = 1.11X – 0.07</td>
<td>0.712 ***</td>
<td>0.761 ***</td>
</tr>
</tbody>
</table>

The correlation coefficients with ** and *** are significant at the 1 and 0.1% probability levels, respectively. The WS and DS indicate wet and dry seasons, respectively.
yielding lines with high protein content from the interspecific cross (WAB56-104/CG 14), we attempted to extrapolate the relationship between paddy yield and protein content per unit area (protein yield) by modifying Figure 2. Figure 3 depicts the relationship between paddy yield and protein yield, which is the total amount of protein per unit land area. The regression lines are adjusted to pass through the origin since 0 protein yield is expected at 0 paddy yield. In all years, a linear regression is strongly applicable to represent that relationship in the interspecific progenies (Fig. 3), suggesting that protein content was not different between low and high yielding interspecific progenies. Since the increase of protein yield with that of paddy yield seems not to reach the ceiling (Fig. 5), the relationship between them in the range of the current paddy yield level will be possibly extrapolated for 1 or 2 t ha⁻¹ higher paddy yield levels than the current highest yielder. Therefore, if interspecific progenies with such a paddy yield level are developed, they could be expected to have a protein content similar to that of the existing ones.

Fig. 2. Relationship between paddy yield and protein content of grain in the interspecific progenies in three different seasons. Solid lines indicate values of the mid-parents for both paddy yield and protein content. The OGP, OSP, B189, WS and DS indicate the *Oryza glaberrima* parent (CG 14), the *O. sativa* parent (WAB56-104), Bouake 189, wet and dry seasons, respectively.

Fig. 3. Relationship between paddy yield and protein yield in the interspecific progenies in three different seasons. Each regression line is adjusted to pass through the origin. Dash lines indicate the 5%-confidence interval for the linear regression line determined for the interspecific progenies. The WS and DS indicate wet and dry seasons, respectively. The correlation coefficient with *** is significant at the 0.1% probability level.
However, it is uncertain if interspecific lines with a protein content higher than the average of the existing interspecific progenies can be developed.

The difference in protein content between WAB450-B-1A1.1, one of the high protein interspecific progenies, and WAB56-104 was 1.32% on the average of the two years excluding 1998 WS when severe zinc deficiency damage was observed. This increase corresponds to about 14% of the protein content of WAB56-104. When the protein content of WAB450-B-1A1.1 is compared to that of Bouake 189, the increment is 2.38% and this value is about 29% of the protein content of Bouake 189. Since several O. glaberrima lines whose protein content was above 13% were identified in our screening, it will be possible to develop new interspecific lines with higher protein content than existing ones if such high-protein O. glaberrima lines are used for the interspecific breeding.

Acknowledgment

The authors would like to express their sincere thanks to Dr. Ryuichi Ishii, Professor of Nihon University, for his valuable suggestions and discussions to initiate this study and produce this paper.

References


* In Japanese with English abstract.
** In French.