Genetic Diversity Studies on Selected Rice (Oryza Sativa L.) Genotypes Based on Amylose Content and Gelatinization Temperature

Article - October 2015

11 authors, including:

- Amos Mawia
  Chinese Academy of Agricultural Sciences
  1 PUBLICATION 0 CITATIONS
  SEE PROFILE

- Mathew Piero Ngugi
  Kenyatta University
  130 PUBLICATIONS 481 CITATIONS
  SEE PROFILE

- Festus Wambua
  Kenyatta University
  24 PUBLICATIONS 60 CITATIONS
  SEE PROFILE

- Daniel Agyirifo
  Kenyatta University
  28 PUBLICATIONS 112 CITATIONS
  SEE PROFILE

Some of the authors of this publication are also working on these related projects:

- Hematological Markers of In Vivo Toxicity View project
- Antipyretic and antinociceptive properties of Harrisonia abyssinica Oliv and Landolphia buchanani(Hallier f) Stapf in animal models View project
Genetic Diversity Studies on Selected Rice (*Oryza Sativa*) Genotypes Based on Amylose Content and Gelatinization Temperature

Mawia A. Musyoki¹, Wambua F. Kioko¹, Ngugi P. Mathew¹, Agyirifo Daniel¹, Karau G. Muriira², Nyamai D. Wavinya¹, Matheri Felix¹, Lagat R. Chemutai¹, Njagi S. Mwenda¹, Mworia J. Kiambi¹ and Ngari L. Ngithi¹

¹Department of Biochemistry and Biotechnology, School of Pure and Applied Sciences, Kenyatta University, P.O. Box 43844-00100, Nairobi, Kenya
²Molecular Laboratory, Kenya Bureau of Standards, Kenya

Abstract

Improving cooking and eating quality of rice is one of the important objectives of many breeding programs. The aim of the study was to carry out genetic diversity studies on selected rice (*Oryza Sativa L.*) genotypes from Kenya and Tanzania based on amylose content and gelatinization temperature using microsatellite markers. Power marker version 3.25 and GenALEX version 6.5 softwares were used to analyze the data. The number of alleles per locus ranged from 2 to 4 alleles with an average of 2.75 alleles across 8 loci obtained in this study. The polymorphic information content (PIC) values ranged from 0.2920 (RM 202) to 0.6841 (RM 141) in all 8 loci with an average of 0.4697. Pair-wise genetic dissimilarity coefficients ranged from 0.9003 to 0.2251 with an average of 0.5627. Maximum genetic similarity was observed between R 2793 and BS 17, Supa and IR64, R 2793 and ITA 310, Soro 5 and ITA 2974. Minimum similarity of was observed between Wahihiwai and BW 196, IR 64 and BW 196. The dendogram based on cluster analysis by microsatellite polymorphism grouped the rice genotypes into 2 clusters effectively differentiating Kenyan and Tanzanian rice genotypes based on amylose content and gelatinization temperature. The results obtained suggested that use of microsatellite markers linked to Quantitative Trait Loci (QTLs) controlling these two traits could effectively be utilized for diversity analysis among diverse rice genotypes.

Keywords: *Oryza sativa; Oryza glaberrima*; Heterozygosity; Germplasm; Genotypes

Introduction

Rice (*Oryza spices*) is a monocotyledonous plant belonging to the family Gramineae and subfamily Oryzoidea. It is cultivated under diverse eco-geographical conditions in various tropical and subtropical countries [1]. Due to its importance as a food crop, rice is being planted on approximately 11% of the Earth’s cultivated land area [2]. It is the grain with the third highest production globally after sugarcane and maize (FAOSTAT, 2012). *Oryza sativa* and *Oryza glaberrima* are the only two cultivated species of rice while the other species are wild. *Oryza sativa* is commonly grown in Asia, North and South America, Europe and Africa. *Oryza glaberrima* is highly grown in West African but due to higher yields of *O.sativa* and *O. glaberrima-sativa* varieties, it is being replaced in most parts of Africa [3].

There are two classes of rice based on starch content, that is, waxy and non waxy rice. Glutinous or waxy rice in which endosperm starch lacks or has very little amylose content consists mainly of amylopectin starch [4]. The ratio of amylose to amylopectin has a major effect on the physical properties of starch. When cooked, the semi-crystalline structure of rice starch is disrupted thus transforming the starch into a softer, edible, and gel-like material [5]. Generally, the amylose content of milled rice is classified into five classes: waxy (0-2%), very low amylose (3-9%), low amylose (10-19%), intermediate amylose (20-24%) and high amylose (above 24%) [6]. The cooking temperature at which water is absorbed and the endosperm starch granule swell irreversibly with subsequent loss of crystalline structure is referred to as gelatinization temperature (GT) [7]. Gelatinization temperature is an important component of rice cooking quality. Rice grain with low gelatinization temperature takes shorter cooking times leading to significant potential savings in fuel costs [8]. Three classes of GT are recognized in rice breeding programs: high (>74°C), intermediate (70-74°C), and low (<70°C) [9,10]. Besides waxy and alk genes that controls amylose content and gelatinization temperature in rice, there are other several QTLs within the rice genome that are linked to these two genes. There is a wide range of rice varieties grown both in Kenya and Tanzania. These rice cultivars are either local landraces or improved varieties and they express different levels of amylose and amylopectin that influences amylose content and gelatinization temperature in rice respectively. Since these two traits are key determinant in cooking and eating qualities of rice, unscrupulous traders often blend rice grains which have good cooking and eating quality traits with grains which have poor cooking and eating quality traits based on amylose content and gelatinization temperature to make more profit from their trade. This causes a negative impact on rice trade and consumption. Accurate evaluation of these two traits is difficult and has hindered development of better varieties with good eating and cooking qualities by rice breeders both in Kenya and Tanzania. The various physicochemical methods commonly used to determine amylose content and gelatinization temperature in rice are often inaccurate and time consuming. However, genetic diversity analysis on these selected rice genotypes from Kenya and Tanzania based on amylose content and gelatinization temperature using microsatellite markers has not yet been studied.

Molecular markers can have a number of applications in agriculture, and their application in rice improvement has been reviewed [11]. Simple Sequence Repeat (SSR) markers are easily available for any...
region of the genome. In rice, SSR markers have been effectively utilized for many purposes such as genetic diversity and relatedness [12], QTL mapping [13], mutation analysis [14] and marker assisted selection [15]. This study was carried out to estimate the pattern and level of genetic diversity and relatedness among selected rice genotypes from Kenya and Tanzania along with other germplasm from Philippine based on amylose content and gelatinization temperature.

Materials and Methods

Plant materials

A total of 13 rice genotypes comprising of local landraces and improved rice genotypes were collected from Mwea Irrigation Agricultural Development Centre (MIAD) in Mwea, Kenya and Kilimanjaro Agricultural Research Institute in Moshi, Tanzania. The name, country of origin and category of the rice genotype chosen for the study are given in (Table 1). Genotype IR 64 was used as the check variety since it is known to have both intermediate amylose content and gelatinization temperature [16].

DNA extraction and SSR marker analysis

Genomic DNA was extracted from seed samples using a modified CTAB method [17]. The quality of the genomic DNA was determined in a 1% agarose gel in 100ml TBE electrophoresis by running 10 μl of genomic DNA at a voltage of 75 for 45 minutes. DNA purity for each sample solution was evaluated using a spectrophotometer which employed the Thermo Scientific Nano drop 2000 system (Wilmington, USA). A set of 8 microsatellite markers shown in (Table 3) covering different genomic regions of rice were selected from published database search for rice Simple Sequence Repeats (SSR) markers [18,19].

The PCR reactions were carried out in Thermal cycler (Bio Rad Inc. USA) with the total reaction volume of 25μl containing, 5 μl of genomic DNA, 1X assay buffer, 200 μM of dNTPs, 2 μM MgCl2, 0.2μM of forward and reverse primer and 1 unit of Taq DNA polymerase (Fermentas Life Sciences). The PCR cycles were programmed as 95oC for 2 min, 94oC for 35 cycles and 72oC for 2 min for each cycle, followed by a final extension of 5 min at 72°C. The amplified products were separated on 1.0 percent agarose gel prepared in 0.5X TBE buffer prepared using Fermenitas Life Sciences. The PCR products were sized against 100bp DNA ladder (Life sciences-USA). The PCR products were visualized under UV-trans illuminator and photographs were taken using gel documentation instrument. The PCR products were sized against 100bp DNA ladder (Life sciences-USA). Clearly resolved, unambiguous bands were scored visually for their presence or absence with each primer. The scores were obtained in the form of matrix with ‘1’ and ‘0’, indicating the presence and absence of bands in each genotype respectively.

Data management and analysis

Using the Power Marker version 3.25 software package [20], the diversity of each accession was analysed on the basis of three statistical parameters: allele number, gene diversity and polymorphism information content (PIC), which measures the genic diversity [21]. Genetic distance was calculated using “C.S Cord 1967” distance [22] followed by phylogeny reconstruction using rooted UPGMA. It was chosen to complement the UPGMA cluster analysis. Furthermore, to reveal the partition and variation within and among the genotypes, analysis of molecular variance (AMOVA) was carried out using GenALEx 6.5 statistical software [23].

Results

Assessment of polymorphism from SSR profiles

Out of 12 SSR markers used, only 8 markers were polymorphic and showed consistent banding patterns and amplification of each genotype and were ultimately chosen for assessing genetic diversity among the rice genotypes studied. A total of 22 alleles were detected at the loci of 8 microsatellite markers across 13 rice genotypes. The allelic richness per locus varied from 2 to 4, with an average of 2.75 alleles. RM 141 produced the highest number of polymorphic alleles while RM 125, RM 202, and RM 253 produced the least number of polymorphic alleles as shown in (Table 2).

Occurrence of null allele was also observed among some genotypes whenever an amplification product could not be detected in their combination. Experiments detecting null alleles were all repeated at least once to ensure that the absence of an amplified product was not as a result of experimental error. Nineteen SSR loci showed null alleles in two to eight of the 13 rice genotypes. The genotypes having the largest proportion of null alleles were Kahogo and IR 54 (null alleles at four loci) and IR 64 and Supa (null alleles at two loci). The frequency of null allele was not included in the genetic diversity calculations for each SSR locus since they might decrease the apparent heterozygosity in a population leading to deviation of genotypes in a sample from Hardy-Weinberg expectations.

Since gene diversity is the measure of expected heterozygosity, the average gene diversity among the 13 rice populations was 0.5503. The gene diversity values ranged from 0.3550 (RM 202) to 0.7337 (RM 141) as shown in Table 3 Polymorphism information content (PIC) value is a measure of polymorphism among varieties for a marker locus used in linkage analysis. The eight SSR markers used in this study produced polymorphic bands as shown in (Figure 1). The PIC value of each marker, which can be evaluated on the basis of its alleles, varied greatly for all tested SSR loci. The values ranged from 0.2920 (RM 202) to 0.6841 (RM 141) with an average of 0.4697 per locus as shown in (Table 2).

Pairwise genetic dissimilarity

A dissimilarity matrix based on the “C.S Cord 1967” shared SSR alleles was used to determine the level of relatedness among the rice genotypes based on amylose content and gelatinization temperature. The pairwise genetic dissimilarity values as shown in table 3 ranged

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Origin</th>
<th>Category of rice</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R 2793</td>
<td>Kenya</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>BS 217</td>
<td>Kenya</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>BS 370</td>
<td>Kenya</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>BW 196</td>
<td>Kenya</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>ITA 310</td>
<td>Kenya</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>SARO 5</td>
<td>Tanzania</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>IR 64</td>
<td>Philippine</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>KILOMERO</td>
<td>Tanzania</td>
<td>Local land race</td>
<td></td>
</tr>
<tr>
<td>RED AFRA</td>
<td>Tanzania</td>
<td>Local land race</td>
<td></td>
</tr>
<tr>
<td>KAHOGO</td>
<td>Tanzania</td>
<td>Local land race</td>
<td></td>
</tr>
<tr>
<td>SUPA</td>
<td>Tanzania</td>
<td>Local land race</td>
<td></td>
</tr>
<tr>
<td>IR 54</td>
<td>Philippine</td>
<td>Improved variety</td>
<td></td>
</tr>
<tr>
<td>WAHIWAHI</td>
<td>Tanzania</td>
<td>Local land race</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Names, origin and category of the rice genotypes used in this study.

from 0.9003 to 0.2251 with an average of 0.5627. Wahiwahi and BW 196; and IR 64 and BW 196 exhibited the highest genetic dissimilarity of 90.03% followed by IR 64 and BS 370; Wahiwahi and BS 217; and Kahogo and ITA 310 with dissimilarity of 78.78%. On the other hand, R 2793 and BS 217; Supa and IR 64; R 2793 and ITA 310; Sario 5 and ITA 310; and Sario 5 and R 2794 depicted the lowest genetic dissimilarity of 22.51%.

Clustering of rice genotypes

A rooted UPGMA tree presented in (Figure 2) was constructed using the “C.S Chord 1967” distance values (Cavalli-Sforza and Edwards, 1967) in Power Marker with tree viewed using Treeview software. It revealed genetic relatedness among the 13 rice genotypes based on amylose content and gelatinization temperature using the 8 microsatellite markers. Genotypes that are derivatives of genetically similar types clustered together. The rice genotypes were clustered into two major groups; that is: group I and group II as shown in Figure 2. Group I consisted of the 5 improved genotypes form Kenya and one improved genotype from Tanzania. Group I was further divided into two subgroups; IA and IB. IA was further divided into two small groups; IA1 and IA2. Group IA1 comprised of R 2793 and BS 217 while group IA2 comprised of Sario 5 and ITA 310. Group IB comprised only of BW 196 and BS 370.

On the other hand, Group II was comparatively diverse since it consisted of 7 genotypes; 5 local landraces from Tanzania and 2 improved genotypes from Philippine. Group II was divided into two major subgroups; IIA and IIB. Major subgroup IIA was further divided into minor subgroups; IIA1 and IIA2. Group IIA1 consisted of Kahogo while subgroup IIA2 consisted of Supa and IR 54. The other major subgroup IIB was further divided into two minor subgroups; IIB1 and IIB2. Subgroup IIB1 consisted only of Red Afaa while subgroup IIB2 comprised of Kilombero, IR 64 and Wahiwahi.

**Principle coordinate analysis (PCoA)**

Principle coordinate analysis (PCoA) based on Nei’s genetic distance (Nei 1972) was used to visualize the genetic relationship among the accessions as shown in (Figure 3). The PCoA supports the results obtained from UPGMA cluster analysis. The first two principle axes accounted for 28.42% and 25.58% variation respectively. The first principle axis comprised of Sario 5, R 2793, ITA 310, and BS 217. The second quadrant comprised of Kilombero, Wahiwahi and IR 64. The third quadrant comprised of Red Afaa, Kahogo and Supa while the forth quadrant comprised of IR 54, BS 370 and BW 196.

**Analysis of molecular variance (AMOVA)**

Analysis of molecular variance (AMOVA) was used to determine the proportion of genetic variation partitioned among and within the 13 rice genotypes. Twenty four percent (24%) (P<0.001) of genetic variation partitioned among genotypes and 76% (P<0.001) within the genotypes as shown in (Table 4).

**Discussion**

Assessment of genetic diversity is a key factor for germplasm conservation, characterization and breeding. Classical breeding affects genetic diversity by selection of combination of outcomes from diverse allele frequencies and leads to favorable effects and loss of diversity [24]. Little was known about genetic diversity among Kenyan and Tanzanian rice genotypes. Additional studies are needed to further understand how these genotypes are adapted to their respective environments.

---

**Table 3:** SC. Cord coefficients of dissimilarity among pairs of 13 rice genotypes.
occurrence was observed by Lapitan VC et al. [28].

targeted gene sequence using the SSR primer pairs. Similar null allele
within a particular gene locus. As a result, it was difficult to amplify the
rice genotypes could have been attributed by chromosomal mutation
lacks that gene’s normal function. Occurrence of null alleles in some
scorable alleles.

used by these authors and selection of SSR markers with different
in average allele per locus might be due to diverse nature of genotypes
rice varieties compared to those from Philippine. In the present study,
a total of 8 SSR markers were used to assess genetic diversity of 13
rice genotypes. The results indicated a considerable level of genetic
diversity among the genotypes used. Different numbers of alleles per
locus were detected in all the 8 SSR markers used. This could have been
attributed by remnant heterozygosity in some varieties and expected
varietal heterogeneity where landrace varieties consist of mixtures
of pure lines that contribute to their broad adaptation in traditional
farming systems. These results were similar to those reported by Shahid
MS et al. [25] using a different group of rice genotypes. In contrast,
the average number of alleles per locus obtained in the present study was
smaller than that reported in previous studies [26,27]. This difference in
average allele per locus might be due to diverse nature of genotypes
used by these authors and selection of SSR markers with different
scorable alleles.

A null allele is a mutant copy of a gene at a locus that completely
lacks that gene’s normal function. Occurrence of null alleles in some
rice genotypes could have been attributed by chromosomal mutation
within a particular gene locus. As a result, it was difficult to amplify the
targeted gene sequence using the SSR primer pairs. Similar null allele
occurrence was observed by Lapitan VC et al. [28].

Gene diversity is the measure of expected heterozygosity. The
genetic diversity mean value of 0.5503 depicted relative heterozygosity
based on amylose content and gelatinization temperature among the
13 rice genotypes studied. The average gene diversity value was lower
than what was reported by Lapitan VC et al. [28] whom obtained a
value of 0.71 but slightly higher than the one obtained by Shahid MS
et al. [25]. The high mean gene diversity value reported by Lapitan
VC et al. [28]. Could be as a result of high rate of exchange of genetic
materials among the rice genotypes studied mostly during their genetic
improvement. The relative heterozygosity reported in our study could
have been attributed by mixing or exchange of genetic materials from
different parental lines especially during rice improvement strategies.
Conventional breeding of rice by rice breeders is one of major
contributory factor to the exchange of genetic material between two
lines resulting into a hybrid variety.

The level of polymorphism was determined by calculating
polymorphism information content (PIC). Polymorphism information
content is a measure of allele diversity at a locus. Loci which are highly
informative have PIC value > 0.5), reasonably informative loci have PIC
values between 0.25 and 0.5, and slightly informative loci have PIC
value <0.25. Out of the 8 SSR markers used in this study, only 3 markers
(RM 141, RM 225, and RM 434) had a PIC value greater than 0.5.
These markers appeared to be highly informative and could therefore
be utilized in marker assisted selection of rice genotypes because they
are capable of distinguishing between genotypes. The mean PIC value
observed in this study was higher than the PIC value of 0.31 recorded by
Sivaranjani AKP et al. [29]. This indicated that the genotypes used in
the present study were more diverse due to differences in origin.

Maximum genetic distance values observed in our study between
Wahiwahi and BW 196; IR 64 and BW 196 indicated high genetic
dissimilarity between them and they showed more divergence.
Chromosomal mutation and diverse geographical origin could be
the contributory factor for the observed genetic dissimilarity between
these rice genotypes. The minimum mean genetic distance values observed
between R 2793 and BS 217; Sora and IR 64; R 2793 and ITA 310; Sora
5 and ITA 310; Sora 5 and R 2794 portrayed less genetic dissimilarity
and having a very close relationship. The observed genetic dissimilarity
is an indication of a common ancestral origin, or perhaps high rate
interbreeding which results to sharing of similar alleles in their genome.
This result was comparable with what was reported by Islam MM et al.
[30] But slightly higher than what was observed by Shahid MS et al.
[25]. A different value of average genetic similarity of 0.79 between 40
rice genotypes was reported by Ravi M et al. [31]. This genetic difference

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Df</th>
<th>SS</th>
<th>MSD</th>
<th>% variation</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among genotypes</td>
<td>1</td>
<td>131.1</td>
<td>65.6</td>
<td>24%</td>
<td>0.001</td>
</tr>
<tr>
<td>Within genotypes</td>
<td>12</td>
<td>1113.8</td>
<td>48.4</td>
<td>76%</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 4: Summary of analysis of molecular variance (AMOVA). Degrees of
freedom (df), sum of squares (SS), mean of square deviation (MSD), % variation
and P-values.
could be due to different group of rice genotypes used. The high level of similarity recorded by this author could be due to the intra specific variation in the germplasm used.

On comparing Kenyan and Tanzanian rice genotypes, it was found that Kenyan genotypes were closely related compared to Tanzanian genotypes, with a mean genetic dissimilarity coefficient of 0.3939 against that of 0.5064 for the Tanzanian genotypes. The genetic closeness of Kenyan rice varieties could be as a result of high intra specific variation, evolution from a common ancestry and introgression of similar traits during the time of genetic improvement. On the other hand, the relatively high genetic dissimilarity witnessed among the Tanzanian varieties could be as a result of having diverse ancestral origins, high gene flow caused by cross pollination among these varieties and chromosomal mutations in their genome.

Clustering of these genotypes together, for example, BW 196 and BS 370 could be as a result of sharing common ancestry or similar genes that grouped together were interpreted to have similar characteristics. The UPGMA cluster analysis by visualizing the relationship between the sample points in a three dimensional space. It was chosen to complement the UPGMA cluster analysis by visualizing the relationship between the sample points in a three dimensional space.

**Supa**, a local landrace from Tanzania and IR 54, an improved cultivar with low amylose content from Philippine clustered together. Based on these results, these two genotypes share common alleles for waxy gene responsible for high amylose content and alk alleles associated with low gelatinization temperature. Therefore, **Supa** genotype does not have good cooking and eating quality characteristics. Genes expressing good quality traits should be introgressed into these genotypes. **Kilombero**, **Wahiwahi** and IR 64 were clustered together. Factors such as sharing of common ancestry and gene flow caused by interspecific gene transfer could be reason behind clustering together. Clustering of **Kilombero** and **Wahiwahi** together with IR 64 is an indication that these two Tanzanian local landraces have good cooking and eating qualities like those of IR 64.

**Conclusion**

The present study provided an overview of genetic diversity based on amylose content and gelatinization temperature among the rice genotypes studied. The UPGMA cluster analysis showed that all 13 rice genotypes could be easily distinguished based on the information generated by the 8 polymorphic SSR markers. The PIC values revealed that RM 141, RM 225, and RM 434 might be the best markers for identification and diversity estimation of rice genotypes. The genetic distance revealed that the Kenyan genotypes had relatively narrow genetic base compared to the Tanzanian genotypes. Therefore, it is highly important not only to conserve these genotypes, but also to reveal their gene pool and unlock other valuable genes for breeding purposes.

**References**


