Electronic Journal of Biotechnology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

Research article

Molecular characterization and genetic diversity of different genotypes of *Oryza sativa* and *Oryza glaberrima*

Caijin Chen [a,1], Wenchuang He [a,1], Tondi Yacouba Nassirou [a], Athanase Nsabiyumva [a], Xilong Dong [a], Yawo Mawunyo Nevame Adedze [b], Deming Jin [a,*]

[a] MOA Key Laboratory of Crop Ecophysiology and Farming System in the Middle Reaches of the Yangtze River, College of Plant Science and Technology, Huazhong Agricultural University, 430070 Wuhan, China
[b] State Key Laboratory of Rice Biology, China National Rice Research Institute, 310006 Hangzhou, China

ARTICLE INFO

Article history:
Received 16 February 2017
Accepted 9 August 2017
Available online xxxx

Keywords:
African rice
Asian rice
Fingerprinting
Food security
Genetic relationship
Microsatellite markers
Molecular profiling
Phylogenetic tree
Polymorphic alleles
Rice breeding
SSR

ABSTRACT

Background: Availability of related rice species is critical for rice breeding and improvement. Two distinct species of domesticated rice exist in the genus Oryza: *Oryza sativa* (Asian rice) and *Oryza glaberrima* (African rice). New rice for Africa (NERICA) is derived from interspecific crosses between these two species. Molecular profiling of these germplasms is important for both genetics and breeding studies. We used 30 polymorphic SSR markers to assess the genetic diversity and molecular fingerprints of 53 rice genotypes of *O. sativa*, *O. glaberrima*, and NERICA.

Results: In total, 180 alleles were detected. Average polymorphism information content and Shannon's information index were 0.638 and 1.390, respectively. Population structure and neighbor-joining phylogenetic tree revealed that 53 genotypes grouped into three distinct subpopulations conforming to the original three groups, except three varieties (IR66417, WAB450-4, MZCD74), and that NERICA showed a smaller genetic distance from *O. sativa* genotypes (0.774) than from *O. glaberrima* genotypes (0.889). A molecular fingerprint map of the 53 accessions was constructed with a novel encoding method based on the SSR polymorphic alleles. Ten specific SSR markers displayed different allelic profiles between the *O. glaberrima* and *O. sativa* genotypes.

Conclusions: Genetic diversity studies revealed that 50 rice types were clustered into different subpopulations whereas three genotypes were admixtures. Molecular fingerprinting and 10 specific markers were obtained to identify the 53 rice genotypes. These results can facilitate the potential utilization of sibling species in rice breeding and molecular classification of *O. sativa* and *O. glaberrima* germplasms.

1. Introduction

Rice is one of the most important crops in the world; and the *Oryza* germplasm serves as the essential resource for rice breeding and contributes significantly to global food security [1]. There are two distinct cultivated species in the genus *Oryza* that both belong to the AA genome with diploid level (2n = 24), *O. sativa* (Asian rice) and *O. glaberrima* (African rice). They were independently domesticated from divergent progenitor wild species in different geographic locations, South Asia and West Africa, respectively [2]. The *O. glaberrima* has many useful traits such as resistance to pests and diseases, and drought tolerance but has lower grain yield potential than *O. sativa* [3,4,5]. Recently, breeders have turned their attention to select elite Asian rice parents because the widespread adoption of similar but improved varieties has decreased genetic diversity of rice gene pool. This might be a major contributing factor for the increased vulnerability to various biotic/abiotic stresses and the yield plateau witnessed in rice production [6]. To overcome these challenges, introduction of new favorable genetic material from a closely related species within the genus *Oryza* is considered a promising approach [7]. African rice could be an important donor germplasm to enrich the rice genetic pool [8,9]. By overcoming the reproductive barriers that existed between the two species, introgression lines such as new rice for Africa (NERICA) have been developed to combine the superior traits of *O. glaberrima* and *O. sativa,* and bridge the genetic gap between the two distinct species [10,11]. Furthermore, utilization of interspecific heterosis *via* partial interspecific hybrid rice between *O.*

* Corresponding author.
E-mail address: djin@mail.hzau.edu.cn (D. Jin).
[1] These authors contributed equally to this work.
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

http://dx.doi.org/10.1016/j.ejbt.2017.08.001

Please cite this article as: Chen C, et al, Molecular characterization and genetic diversity of different genotypes of *Oryza sativa* and *Oryza glaberrima*, (2017), http://dx.doi.org/10.1016/j.ejbt.2017.08.001

sativa and introgression lines carrying *O. glaberrima* genes could be a promising novel approach to raise the grain yield potential [12].

Several previous researches have revealed the genetic diversity of different rice germplasms. An assessment of the molecular diversity of 79 *O. glaberrima* germplasms from Mali (West Africa) revealed that the populations from different locations were highly differentiated [13]. Genetic diversity and population structure were investigated using 93 simple sequence repeats (SSR) markers for 198 accessions of *O. glaberrima* collected from 12 different countries in West Africa. Genetic evidence indicated that 67% of *O. glaberrima* accessions carry some level of admixture with *O. sativa* and that natural interspecific outcrossing might occur because *O. glaberrima* was often grown in combination with *O. sativa* in West Africa [14]. Similar results were reported by Barry et al. [15] who found a close genetic relationship between *O. sativa* and some *O. glaberrima* accessions using 11 SSR markers and 26 morpho-physiological descriptors [15]. Similarly, reports also describe the genetic diversity of *O. sativa*. Two subgroups including *indica* and *japonica* as well as six sub-subgroups were found within a primary *O. sativa* core collection [16]. Eight subpopulations were found to correspond to major geographic regions among 103 *O. sativa* accessions studied [17]. However, to our knowledge, the comparison of genetic diversity of different cultivated species of rice genotypes has not been reported. A better understanding of the extent and distribution of genetic diversity within and between different genotype groups is essential not only to assist plant breeders in the selection of parents but also to provide a more rational basis for expanding the gene pool and for identifying materials that harbor alleles valuable for plant improvement.

Based on the phenotypic and genotypic data, some rice core collections were evaluated for germplasm with similar genetic diversity as the entire rice collection; this can be an effective and convenient tool for rice breeders. However, it is difficult to unambiguously identify cultivars from other groups using conventional morphological characteristics owing to effects of environmental factors [18]. Fingerprinting with molecular markers allows a precise, objective, and rapid cultivar identification, which has been proven to be an efficient tool for crop germplasm characterization, collection, and management [19]. A variety of molecular markers can be used to evaluate the genetic diversity and establish the fingerprint of rice genotypes, such as SSRs and single nucleotide polymorphisms (SNPs). Until now, SSR markers have been widely used for assessment of genetic diversity and establishment of unique fingerprint owing to their abundance, codominant inheritance, high polymorphism, reproducibility, ease of assay by polymerase chain reaction (PCR), and relatively low cost [20,21]. Recently, SNPs have received increased attention because they occur at a much higher frequency in the genome than SSRs. However, most SNPs are biallelic; thus, an SNP marker has less information content than an SSR marker [22]. Furthermore, SSR markers have their own advantages as compared to SNP markers for population genetics analysis [23] and are still used widely in the construction of molecular fingerprinting databases. A molecular fingerprinting database of 49 rice cultivars was constructed using 24 SSR markers [24]. Unique DNA profiles of conventional *japonica* rice from Taihu Lake area were established based on 24 SSR markers [25]. SSR markers have been used in many crop species for cultivar identification, such as wheat [26], maize [27], bean [28], tomato [29,30], and other crops. Additionally, the fingerprint can be used for the protection of plant genetic resources.

The objectives of this study were to (1) investigate the genetic diversity, variability and molecular phylogeny; and (2) establish an effective and low-cost encoding method for the molecular fingerprinting of 53 rice genotypes including *O. glaberrima, O. sativa*, and NERICA, using SSR markers. This study will enhance our understanding of genetic diversity among different genotypes of *O. sativa* and *O. glaberrima* and facilitate the use of diverse germplasm in rice breeding.

2. Materials and methods

2.1. Experiment materials

A total of 53 rice genotypes were collected for this study, including 18 African rice accessions, 10 NERICA varieties, 23 Asian rice varieties, and two tropical weedy rice accessions that were collected from six different countries (Table S1). Twenty-five *O. sativa* genotypes in the germplasm pools, as well as their known ancestors, were distributed in nine Southeast Asian countries such as China, India, and Indonesia (Fig. 1; Table S1).

2.2. Pedigree analysis

The origin and pedigree information of rice genotypes were investigated using the International Rice Information System database (http://irri.org/tools-and-databases/international-rice-information-system) and the China Rice Data Center (http://www.ricedata.cn/variety/) resources. Pedigree plotting was performed using Pedigraph 2.2 software [31].

2.3. DNA extraction and SSR analysis

Genomic DNA was extracted from young and healthy leaves of plants, using the cetyltrimethylammonium bromide method [32] with minor modifications. Quality of the isolated DNA was checked on 1% agarose gel and quantity was determined using ND-1000 spectrophotometer (NanoDrop Technologies, USA).

A total of 118 SSR primers were used to evaluate the polymorphism of 12 genotypes comprising four *O. glaberrima*, four *O. sativa*, and four NERICA genotypes. To screen the 53 genotypes, 30 SSR primers were selected based on their polymorphism and distribution on the chromosomes (Table S2).

The polymerase chain reaction (PCR) was performed in 20 μL volume, containing 2 μL (50 ng μL^{-1}) DNA, 0.2 μL (5 U μL^{-1}) *Taq* polymerase, 2 μL 10X PCR buffer, 1.2 μL (10 mM) $MgCl_2$, 0.4 μL dNTP, 2 μL primer pairs, and 12.2 μL sterile double distilled water. The PCR cycling parameters were as follows: initial preheating and initial denaturation at 94°C for 4 min, followed by 10 cycles of denaturing at 94°C for 30 s, annealing at 55°C to 65°C for 30 s, decreasing 1°C per cycle, and extension at 72°C for 45 s. The last 22 cycles were at 94°C for 30 s, 55°C for 30 s and 72°C for 45 s, followed by a final extension at 72°C for 10 min. The SSR products were separated on a denaturing 6% polyacrylamide gel and visualized by silver staining. The electrophoretic bands were scored as 1 (present) or 0 (absent) to form a raw data matrix for further analysis.

2.4. Data analysis

POPGENE 3.2 software was used to calculate the number of effective alleles [33], expected heterozygosity [34], Shannon's information index [35], and allele frequency. Allelic polymorphic information content (PIC) was calculated using the following formula:

$$PIC = 1 - \sum_{i=1}^{n} P_i^2$$

where n is the total number of alleles detected for a given marker locus and P_i is the frequency of the ith allele in the set of genotypes investigated [36].

The genetic distance (GD) was calculated using the following formula:

$$GD = 1 - GS$$

where GS was calculated using the NTSYS-pc ver. 2.1e software with option of DICE coefficient.

Please cite this article as: Chen C, et al, Molecular characterization and genetic diversity of different genotypes of *Oryza sativa* and *Oryza glaberrima*, (2017), http://dx.doi.org/10.1016/j.ejbt.2017.08.001