Available online at www.sciencedirect.com
ScienceDirect
J. Genet. Genomics 36 (2009) 435–442
JOURNAL OF GENETICS AND GENOMICS
www.jgenetgenomics.org

# Subspecies-specific intron length polymorphism markers reveal clear genetic differentiation in common wild rice (*Oryza rufipogon* L.) in relation to the domestication of cultivated rice (*O. sativa* L.)

Xiangqian Zhao [a], Long Yang [a], Yan Zheng [a], Zhaohua Xu [a], Weiren Wu [a, b, *]

[a] *Department of Agronomy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310029, China*
[b] *College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China*

Received for publication 2 November 2008; revised 1 March 2009; accepted 17 March 2009

**Abstract**

It is generally accepted that *Oryza rufipogon* is the progenitor of Asian cultivated rice (*O. sativa*). However, how the two subspecies of *O. sativa* (*indica* and *japonica*) were domesticated has long been debated. To investigate the genetic differentiation in *O. rufipogon* in relation to the domestication of *O. sativa*, we developed 57 subspecies-specific intron length polymorphism (SSILP) markers by comparison between 10 *indica* cultivars and 10 *japonica* cultivars and defined a standard *indica* rice and a standard *japonica* rice based on these SSILP markers. Using these SSILP markers to genotype 73 *O. rufipogon* accessions, we found that the *indica* alleles and *japonica* alleles of the SSILP markers were predominant in the *O. rufipogon* accessions, suggesting that SSILPs were highly conserved during the evolution of *O. sativa*. Cluster analysis based on these markers yielded a dendrogram consisting of two distinct groups: one group (Group I) comprises all the *O. rufipogon* accesions from tropical (South and Southeast) Asia as well as the standard *indica* rice; the other group (Group II) comprises all the *O. rufipogon* accessions from Southern China as well as the standard *japonica* rice. Further analysis showed that the two groups have significantly higher frequencies of *indica* alleles and *japonica* alleles, respectively. These results support the hypothesis that *indica* rice and *japonica* rice were domesticated from the *O. rufipogon* of tropical Asia and from that of Southern China, respectively, and suggest that the *indica-japonica* differentiation should have formed in *O. rufipogon* long before the beginning of domestication. Furthermore, with an *O. glaberrima* accession as an outgroup, it is suggested that the *indica-japonica* differentiation in *O. rufipogon* might occur after its speciation from other AA-genome species.

*Keywords*: *Oryza sativa*; *Oryza rufipogon*; evolution; SSILP marker

## Introduction

Rice is one of the most important staple food crops, which feeds more than half of the world's human population. The genus *Oryza* consists of ~23 species with diploid genomes (AA, BB, CC, EE, FF, GG; 2n = 24) or tetraploid genomes (BBCC, CCDD, HHJJ; 2n = 48) (Vaughan et al., 2003). Seven species carry the AA genome, including two cultivated species, *O. sativa* and *O. glaberrima* (Khush, 1997). *O. sativa* consists of two subspecies: *indica* and *japonica*. The subspecies *japonica* can be further divided into two ecotypes: temperate and tropical. But the term *japonica* usually refers to the temperate ecotype; while the tropical ecotype is often called *javanica*. In this paper, we shall follow this custom unless explained specifically. *O. sativa* is cultivated worldwide but mostly in South and Southeast Asia, while *O. glamberrima* is cultivated only in

* Corresponding author. Tel & Fax: +86-571-8697 1910.
*E-mail address*: wuwr@zju.edu.cn

DOI: 10.1016/S1673-8527(08)60133-2

a limited area in West Africa. Therefore, *O. sativa* is more important in agriculture.

Phylogenetic analysis using molecular markers have suggested that *O. sativa* and *O. glaberrima* were domesticated from the wild rice species *O. rufipogon* and *O. barthii*, respectively (Wang et al., 1992; Bautista et al., 2001; Ishii et al., 2001). According to the difference in life cycle and habitat preference, *O. rufipogon* can be divided into two ecotypes, perennial and annual (Oka and Morishima, 1967; Oka, 1988). As *O. sativa* is an annual species, it has been suggested that the annual ecotype of *O. rufipogon* (also named *O. nivara*) might be the most recent progenitor of *O. sativa* (Khush, 1997). However, there are evidences that both the perennial and annual ecotypes could be possibly the progenitors of *O. sativa* (Chang, 1976; Cheng et al., 2003). Hence, we do not differentiate the two ecotypes of *O. rufipogon* in this paper.

The precise times and locations of domestication of *O. sativa* have been debated. There are two hypotheses. One hypothesis suggests that *O. sativa* has a monophyletic origin (Oka, 1974; Oka and Morishima, 1982): *indica* rice was firstly domesticated from wild rice and then *japonica* rice was derived from *indica* rice as an adaptive type to the high elevation and high latitude (Ting, 1957; Chang, 1976; Oka, 1988). The other hypothesis suggests that *O. sativa* has diphyletic or polyphyletic origins: *indica* rice was domesticated within a region south of the Himalaya mountain range, whereas *japonica* rice was domesticated from wild rice in southern China (Khush, 1997). Most biochemical and molecular evidences appear to support the second hypothesis (Second, 1982; Morishima, 1986; Dally and Second, 1990; Wang et al., 1992; Mochizuki et al., 1993; Hirano et al., 1994; Wang and Sun, 1996; Cheng et al., 2003; Londo et al., 2006; Rakshit et al., 2007). However, the recently cloned domestication gene *sh4*, which is responsible for the reduction of grain shattering from wild rice to cultivated rice, appears to have originated only once, giving support to the monophyletic origin hypothesis (Li et al., 2006).

Molecular markers are useful tools for the analysis of genetic diversity and phylogenetic relationship. Unlike the sequence-based analysis, which usually involves only one or a few loci, marker-based analysis usually involves many loci and therefore could have a better representation for genetic variation. Various molecular markers have been used for the analysis and the most frequently used markers are random amplified polymorphic DNA (RAPD) (Williams, 1990), inter-simple sequence repeat (ISSR) (Zietkiewicz, 1994) and simple sequence repeat (SSR) (Becker and Heun, 1995). Because the loci detected by these markers are mainly located in non-coding regions, these markers are generally suitable for the analysis of genetic diversity within species, but may be too variable for the analysis of phylogenetic relationship between species, especially between distant species. In other words, these markers are not ideal for evolutionary studies.

Recently, a new molecular marker system termed intron length polymorphism (ILP) has been developed in rice by comparing the draft genomic sequences of *indica* cultivar 93-11 and *japonica* cultivar Nipponbare (Wang et al., 2005). ILP is a codominant marker and can be conveniently detected by polymerase chain reaction (PCR) with a pair of primers designed on flanking exons. ILPs are comparable among different species (Wang et al., 2005) because the exon-intron structures of genes are highly conservative across species (Rogozin et al., 2003; Roy and Gilbert 2005; Lin et al., 2006; Yang et al., 2007). In addition, ILPs mainly exist among (sub)species or higher taxonomic ranks and are usually of low frequency within (sub)species. So, the alleles (patterns of electrophoretic bands) of ILP markers are usually (sub)species-specific. It has been found that many ILPs between 93-11 and Nipponbare appear to be subspecies-specific (Wang et al., 2005), suggesting that ILPs were relatively conservative during the domestication of *O. sativa*. It is possible that these subspecies-specific ILP (SSILP) markers could well preserve the genetic information from wild progenitors and therefore could reflect the original genetic differentiation between the ancestors of *indica* rice and *japonica* rice. Moreover, introns usually have no biological functions, although some of them might have the function of regulating gene expression (Rose, 2008). Therefore, introns (and so ILP markers) are usually neutral for selection. Examining these neutral regions would allow for historical inference of the difference between *O. sativa* and *O. rufipogon*, where evolutionary forces such as gene flow and genetic drift predominate (Londo et al., 2006). The above desirable features suggest that ILP markers should be quite suitable for studying rice evolution. In the work described here, we utilized SSILP markers to investigate the genetic differentiation of *O. rufipogon* in relation to the genetic differentiation between the two subspecies of *O. sativa*.

## Materials and methods

### *Plant materials*

Seeds of 103 *O. rufipogon* accessions and 20 *O. sativa*

accessions, including 10 *indica* cultivars (93-11, Aijiaonante, Aizizhan, Guangluai-4, IR8, Jaya, Peta, V20, Xieqingzao, Zhenshan-97) and 10 *japonica* cultivars (Balilla, Bluebonnet, Cripto, Dawn, Dong jin byeo, Jia-59, Koshihikari, Lemont, Nipponbare, Xiushui-11), were kindly provided by the International Rice Research Institute (IRRI) and the China National Rice Research Institute (CNRRI). The rice seeds were germinated and grown in a growth chamber with 16 h of light and 8 h of darkness at 25°C. Any *O. rufipogon* accession that generated at least 10 seedlings was used for the study. In addition, leaves of 14 and 6 *O. rufipogon* accessions originated from the mainland and Hainan province (island) of China were kindly provided by CNRRI and Hainan Institute of Tropical Agricultural Resources (HITAR), respectively.

### *Screening of SSILP markers*

A total of 123 ILP primer pairs showing codominant polymorphisms between *indica* cultivar 93-11 and *japonica* cultivar Nipponbare were used. These ILP markers were previously shown to be potentially subspecies-specific based on tests on three typical *indica* cultivars and three typical *japonica* cultivars (Wang et al., 2005). To obtain stricter SSILP markers, we tested these primer pairs again with the 10 *indica* cultivars and the 10 *japonica* cultivars mentioned above. Any primer pair that generated 93-11's band pattern in all the *indica* cultivars and Nipponbare's band pattern in all the *japonica* cultivars was taken as a SSILP marker and was used for the analysis of wild rice accessions.

### *DNA extraction and PCR*

Genomic DNA of each accession was extracted from leaves using CTAB method (Murray and Thompson, 1980) with modification. PCR was performed in a 15 μL reaction mixture containing 50 ng template DNA, 0.5 μmol/L of each primer, 200 μmol/L of each dNTP, 1.5 mmol/L $MgCl_2$, 0.1% Triton X-100 and 1 U *Taq* polymerase and 1.5 μL of 10 × PCR reaction buffer. The PCR procedure was: 5 min initial denaturation at 94°C; 35 cycles of 30 s denaturation at 94°C, 30 s anneal at 55°C and 1 min extension at 72°C; and 5 min final extension at 72°C. For most primers, PCR products were separated by 6% non- denaturing PAGE (250 V, 2 h) and visualized by silver stain. For some primers, 2% agarose gel was used for separating PCR products.

### *Data analysis*

Every PCR band polymorphic among the accessions analyzed was scored as 1 (present) or 0 (absent). The ratios of shared PCR bands and similarity coefficients between accessions were quantified according to Nei (1978). Based on the PCR data, a dendrogram was constructed using the UPGMA (unweighted pair group method with arithmetic mean) method and the reliability of the dendrogram was assessed using bootstrap method implemented by the program FreeTree (Hampl et al., 2001).

## Results

### *SSILP markers and their conservativity*

Fifty-seven (46.34%) of the 123 ILP markers tested were found to be highly subspecies-specific (Supplemental Table 1). Each of them showed only a single allele (band) in the 10 *indica* cultivars and a different single allele in the 10 *japonica* cultivars. Therefore, they could be taken as SSILP markers and the two different alleles of each marker could be taken as *indica* allele and *japonica* allele, respectively. These SSILP markers were randomly distributed in the rice genome. The number of SSILP markers on each chromosome varied from 2 (on chromosome 11) to 8 (on chromosomes 2 and 6) with an average of 4.75. The genes where the SSILP markers were located involved a broad spectrum of molecular functions including ribosomal protein, protein precursors, reductase, isozyme, synthetase, kinase and so on. Hence, the SSILP markers could well reflect the genetic diversity between *indica* rice and *japonica* rice. Based on these 57 SSILP markers, we defined a standard *indica* rice and a standard *japonica* rice, which contains all the *indica* alleles and all the *japonica* alleles, respectively.

The primers of the SSILP markers screened were used to analyze 73 *O. rufipogon* accessions (Table 1). In addition, an African cultivated rice (*O. glaberrima*) accession GLA-AF-1, which also carries the AA genome, was analyzed with the SSILP markers as an outgroup. A total of 144 polymorphic bands were amplified by the SSILP primers in the *O. rufipogon* accessions with a variation range of 2–6 and an average of 2.53 bands per primer pair. Most (35; 61.4%) of the SSILP primer pairs detected two alleles only (i.e., the *indica* allele and *japonica* allele); 17 (29.82%) primer pairs detected an additional allele apart from the *indica* allele and *japonica* allele; 3 (5.26%)

Table 1
Accessions of *O. rufipogon* used in this study

| No. | Name/Code | Origin | No. | Name/Code | Origin | No. | Name/Code | Origin |
|---|---|---|---|---|---|---|---|---|
| 1 | DX-JX-CN-1 | China (A) | 26 | BD-5 | Bangladesh | 51 | LK-1 | Sri Lanka |
| 2 | DX-JX-CN-2 | China (A) | 27 | BD-6 | Bangladesh | 52 | LK-2 | Sri Lanka |
| 3 | DX-JX-CN-3 | China (A) | 28 | BD-7 | Bangladesh | 53 | MM-1 | Myanmar |
| 4 | DX-JX-CN-4 | China (A) | 29 | BD-8 | Bangladesh | 54 | MM-2 | Myanmar |
| 5 | KP-GD-CN-1 | China (B) | 30 | ID-1 | Indonesia | 55 | MY-1 | Malaysia |
| 6 | KP-GD-CN-2 | China (B) | 31 | ID-2 | Indonesia | 56 | MY-2 | Malaysia |
| 7 | KP-GD-CN-3 | China (B) | 32 | ID-3 | Indonesia | 57 | NP-1 | Nepal |
| 8 | KP-GD-CN-4 | China (B) | 33 | ID-4 | Indonesia | 58 | NP-2 | Nepal |
| 9 | SS-GD-CN-1 | China (C) | 34 | ID-5 | Indonesia | 59 | PG-1 | Papua New Guinea |
| 10 | BS-GX-CN-1 | China (D) | 35 | ID-6 | Indonesia | 60 | PG-2 | Papua New Guinea |
| 11 | BS-GX-CN-2 | China (D) | 36 | IN-1 | India | 61 | PG-3 | Papua New Guinea |
| 12 | BS-GX-CN-3 | China (D) | 37 | IN-2 | India | 62 | PG-4 | Papua New Guinea |
| 13 | BS-GX-CN-4 | China (D) | 38 | IN-3 | India | 63 | TH-1 | Thailand |
| 14 | BS-GX-CN-5 | China (D) | 39 | IN-4 | India | 64 | TH-2 | Thailand |
| 15 | ZP-FJ-CN-1 | China (E) | 40 | IN-5 | India | 65 | TH-3 | Thailand |
| 16 | QH-HN-CN-1 | China (F) | 41 | IN-6 | India | 66 | TH-4 | Thailand |
| 17 | QH-HN-CN-2 | China (F) | 42 | IN-7 | India | 67 | VN-1 | Vietnam |
| 18 | QH-HN-CN-3 | China (F) | 43 | KH-1 | Cambodia | 68 | VN-2 | Vietnam |
| 19 | QH-HN-CN-4 | China (F) | 44 | KH-2 | Cambodia | 69 | VN-3 | Vietnam |
| 20 | QH-HN-CN-5 | China (F) | 45 | KH-3 | Cambodia | 70 | VN-4 | Vietnam |
| 21 | LS-HN-CN-1 | China (G) | 46 | KH-4 | Cambodia | 71 | VN-5 | Vietnam |
| 22 | BD-1 | Bangladesh | 47 | KH-5 | Cambodia | 72 | VN-6 | Vietnam |
| 23 | BD-2 | Bangladesh | 48 | LA-1 | Laos | 73 | VN-7 | Vietnam |
| 24 | BD-3 | Bangladesh | 49 | LA-2 | Laos | | | |
| 25 | BD-4 | Bangladesh | 50 | LA-3 | Laos | | | |

A: Dongxiang, Jiangxi; B: Kaiping, Guangdong; C: Sanshui, Guangdong; D: Baise, Guangxi; E: Zhangpu, Fujian; F: Qionghai, Hainan; G: Lingshui, Hainan.

detected two additional alleles; and one (1.75%) each detected three and four additional alleles, respectively. For most of the markers with more than two alleles, the additional allele(s) did not exist alone but together with the *indica* allele or/and *japonica* allele in an accession. Only five markers (RI2813, RI05249, RI05304, RI05407 and RI05559) did not detect the *indica* alleles and *japonica* alleles but the additional allele(s) in a few accessions (ranging from 1 accession for RI05407 to 26 accessions for RI05249). The results indicated that the *indica* alleles and *japonica* alleles of the SSILP markers were predominant in *O. rufipogon* and were conserved during the domestication of *O. sativa*.

In addition, it is interesting that the alleles of the SSILP markers detected in the *O. glaberrima* accession were also mainly the *indica* alleles and *japonica* alleles and were all included in those found in the *O. rufipogon* accessions with only a few exceptions, implying that the *indica* alleles and *japonica* alleles must have existed before the speciation of *O. rufipogon*. The feature of conservativity of the SSILP markers is desirable for the study of rice evolution.

### *Differentiation among the wild rice accessions*

A dendrogram consisting of the 73 *O. rufipogon* accessions and the *O. glaberrima* accession (GLA-AF-1) as well as the defined standard *indica* rice and *japonica* rice was constructed based on the SSILP data (Fig. 1). The *O. glaberrima* accession was classified as an independent group as expected and the rest accessions were clearly classified into two major groups (Group I and Group II), both with very high bootstrap probabilities. Group I contains all the *O. rufipogon* accessions from tropical (South and Southeast) Asia (including Hainan province of China)

and the standard *indica* rice, while Group II contains all the *O. rufipogon* accessions from Southern Mainland China and the standard *japonica* rice. This result indicates that *O. rufipogon* has been genetically differentiated into two distinct groups corresponding to the two subspecies of Asian cultivated rice: Group I is closer to *indica* rice and Group II is closer to *japonica* rice; and the differentiation has a clear geographic pattern: Group I is distributed in tropical Asia, while Group II is distributed in Southern Mainland China.

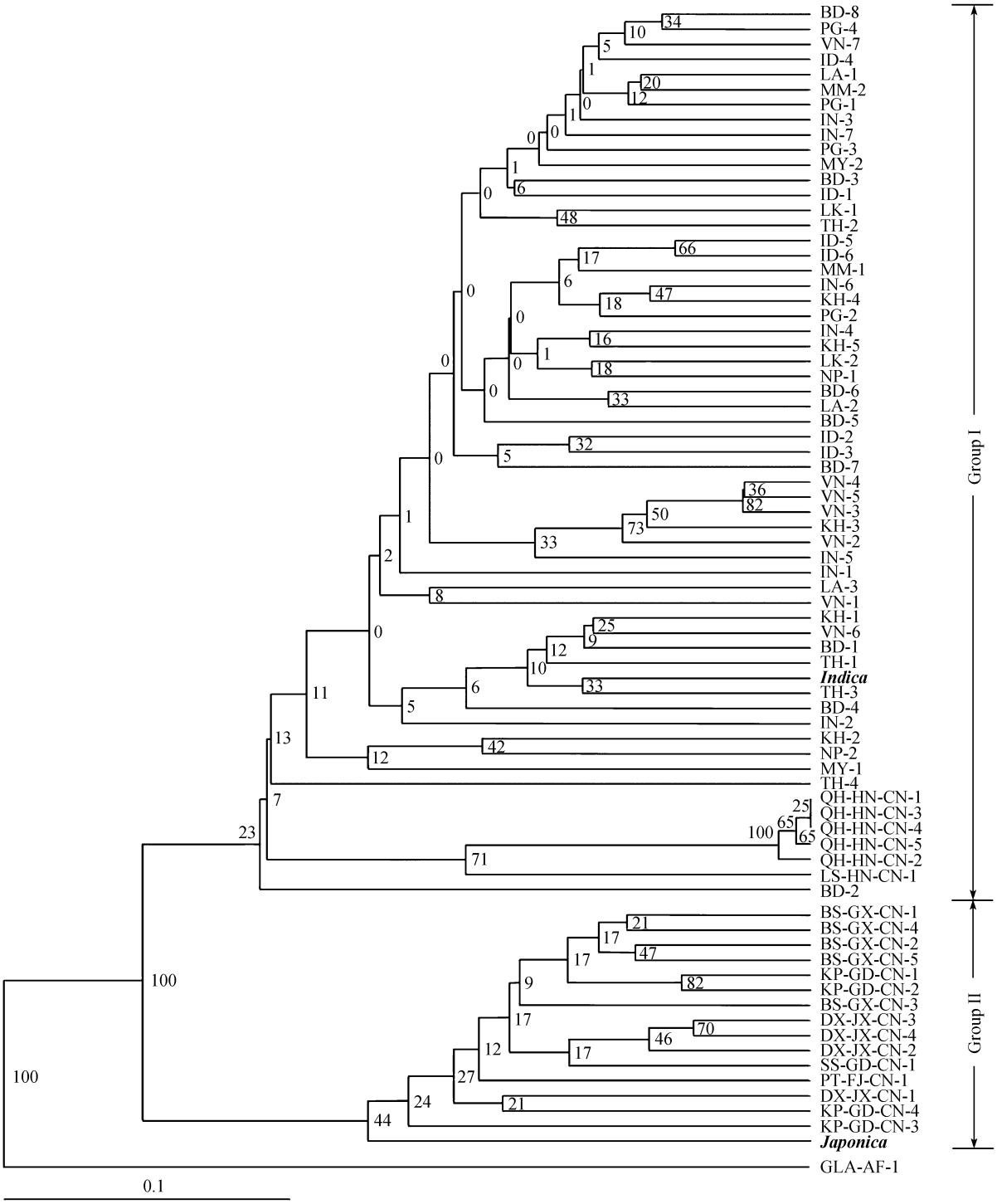


Fig. 1. Dendrogram of 73 *O. rufipogon* accessions and the *O. glaberrima* accession as well as the defined standard *indica* rice and *japonica* rice. The number beside each node is bootstrap probability estimated based on 1,000 times of resampling.

The result of cluster analysis was consistent with the frequency distribution of *indica* allele or *japonica* allele percentage in each *O. rufipogon* accession, which is defined as $n_i/(n_i+n_j)$ or $n_j/(n_i+n_j)$, where $n_i$ and $n_j$ are the numbers of *indica* alleles and *japonica* alleles in the accession, respectively. As the percentages of *indica* alleles and *japonica* alleles are complementary, we need only examine one of them. In the calculation of percentage of *indica* alleles, three accessions from Qionghai (No. 16, 18 and 19 in Table 1) were merged as one accession because they were very similar to each other (Fig. 1). The two groups of *O. rufipogon* accessions showed quite different frequency distributions of percentage of *indica* alleles (Fig. 2). Group I had much higher percentage than Group II on average. This explained why Group I was closer to *indica* rice. This was also coherent with the *indica* allele frequencies of individual markers in the two groups (Fig. 3). The *indica* allele frequencies of most of the markers as well as the average *indica* allele frequency were higher in Group I than in Group II, and the difference of *indica* allele frequency between the two groups was significant at 5% or higher level for more than half of the markers.

## Discussion

We have seen that SSILP markers have two important features. First, they reflect the genetic differentiation between the two subspecies (*indica* and *japonica*) of cultivated rice. Second, the two subspecies-specific alleles (*indica* alleles and *japonica* alleles) are generally conserved and predominant in the wild progenitor species. In addtion, as we have pointed out in the Introduction, SSILP markers are neutral for selection in general since introns are usually functionless. These desirable features of SSILP markers enable us to investigate the genetic differentiation of *O. rufipogon* in terms of the genetic differentiation of *O. sativa*. This study indicates that *O. rufipogon* has been clearly differentiated into two distinct groups parallel to the differentiation of the two subspecies of *O. sativa*. This suggests

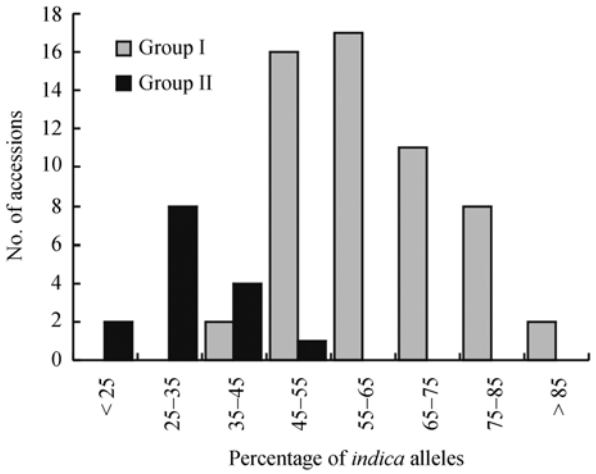


Fig. 2. Frequency distribution of *indica* allele percentage in individual *O. rufipogon* accessions from tropical Asia (Group I) and Southern Mainland China (Group II).

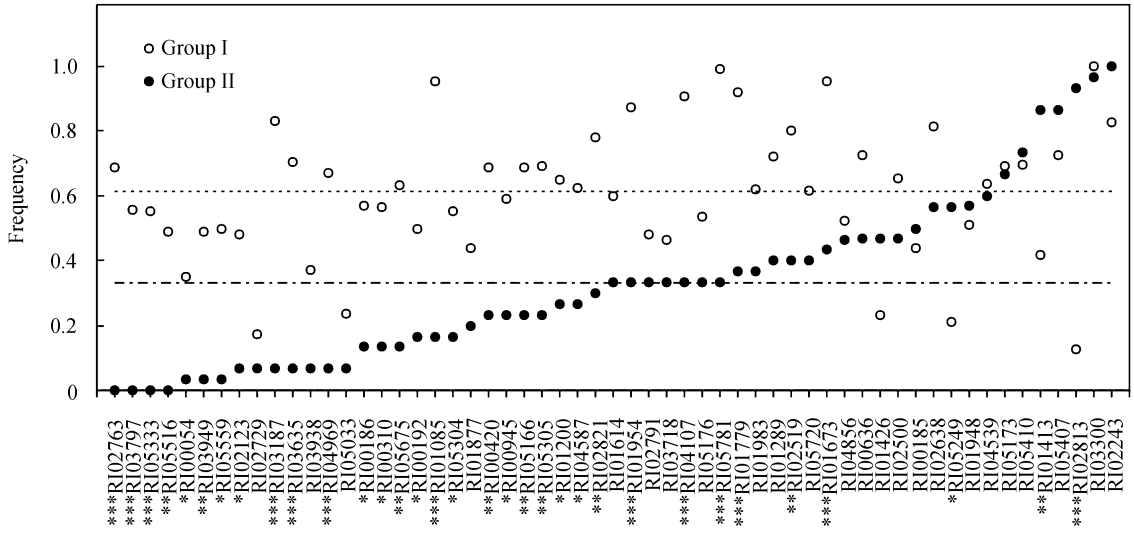


Fig. 3. *Indica* allele frequencies of each SSILP marker in the subpopulations of *O. rufipogon* from tropical Asia (Group I) and Southern Mainland China (Group II). *, ** and ***: significant at 0.05, 0.01 and 0.001 levels in *t*-tests for the difference between Group I and Group II. The horizontal dot line and dash-dot line indicate the means of *indica* allele frequency in Group I and Group II, respectively.

that *indica* rice and *japonica* rice were domesticated from tropical Asia and Southern Mainland China, respectively, giving support to the hypothesis of independent origins of *indica* rice and *japonica* rice (Second, 1982; Ishii et al., 1988; Wang et al., 1992). Previous studies have suggested that the divergence between *indica* rice and *japonica* rice began about 0.4 million years ago (Ma et al., 2004). Therefore, the two distinct groups of *O. rufipogon* might have formed long before the domestication of *O. sativa*. Moreover, this study shows that the outgroup *O. glaberrima* is apparently separated from the group of *O. rufipogon* and *O. sativa* in the dendrogram (Fig. 1), implying that the *indica-japonica* differentiation in *O. rufipogon* might occur after its speciation from other AA-genome species.

A noticeable point revealed in this study is that the differentiation of the two groups in *O. rufipogon* has an apparent geographic pattern—there is a clear geographical borderline between them. The borderline is along the Himalayas, which might serve as a natural barrier to prevent gene communication between the two sides. Hence, geographic isolation must have played an important role in the genetic differentiation between the two groups of *O. rufipogon* (Lu et al., 2002).

In each *O. rufipogon* accession analyzed in this study, most of the SSILP loci appeared to be monomorphic (i.e., only either the *indica* allele or the *japonica* allele existed); only a smaller proportion (20.39% on average, ranging 0–49.12%) of the SSILP loci remained polymorphic (containing both the *indica* allele and the *japonica* allele simultaneously). But almost all of the SSILP markers (with only one exception) showed the *indica* allele vs. *japonica* allele polymorphism in at least one *O. rufipogon* accession each. The result indicates that the *indica* alleles or *japonica* alleles of most of the SSILP loci had been randomly fixed in each *O. rufipogon* accession. The fixation of SSILP alleles might result mainly from genetic drift because SSILP markers are generally neutral for selection. This implies that genetic drift was an important factor for the differentiation between the two *O. rufipogon* groups.

Nevertheless, natural selection might be also a possible cause for the allele fixation of some SSILP markers. The latitude range of Southern Mainland China is higher than that of tropical Asia. Therefore, Southern Mainland China is cooler than tropical Asia. This ecological difference could exert selection pressure on genes related to environmental adaptability, making the two *O. rufipogon* groups adapted to the different ecological conditions. It is known that *indica* rice and *japonica* rice are adapted to warmer and cooler ecological conditions, respectively. This appears to be parallel to the difference between the two *O. rufipogon* groups. Tang et al. (2006) found that different sub-populations of *O. sativa* captured different portions of the genetic diversity of the ancestral *O. rufipogon* population, and much of the diversity (*indica* 9311-like haplotype vs. *japonica* Nipponbare-like haplotype) observed in the high divergence regions, which covered at least 6% of the genome, was preserved. They hypothesized that the two divergent haplotypes in the high divergence regions of the genome were adapted to different geographical and ecological environments, contributing to the enormous range of phenotypic variation observed among *O. sativa* varieties. The finding suggests that the *indica-japonica* differentiation preserved in some genomic regions might be related to geographical and ecological adaptation. By referring to Gramene website, we have found that several SSILP markers investigated in this study are located near to quantitative trait loci (QTLs) conferring cold tolerance. For example, RI02519 on chromosome 1 is close to AQAV003 (QTL accession ID); RI01954 on chromosome 3 is close to AQAV005; and RI02729 on chromosome 6 is close to AQDU002. In addition, several genes in which the SSILP markers are located appear to be related to environmental adaption according to their molecular functions annotated. For example, RI01614 is a UV-damaged DNA-binding protein 1; RI04856 is a heat shock protein; and RI02123 and RI05305 are members of cytochrome P450 family. These SSILP markers could probably be selected if their closely linked QTLs or corresponding genes were subjected to natural selection. Moreover, it is seen in this study that there are several SSILP markers showing zero (or very low) *japonica* allele frequency in Group II but high *japonica* allele frequency in Group I (e.g., RI02763, RI03797 and RI05333) or, contrarily, zero (or very low) *indica* allele frequency in Group I but high *indica* allele frequency in Group II (e.g., RI01085, RI05781 and RI01673) (Fig. 3). Although the phenomenon could result from sampling error, genetic drift or other causes, natural selection could possibly be an important factor. Further investigation on these markers would be helpful for elucidating the issue.

## Acknowledgement

This work was supported by the National Basic Research Program (973 Program) of China (No. 2006CB101708) and the National Hi-Tech Research and

Development Program (863 Program) of China (No. 2006AA10Z1E2).

## Supplemental data

Supplemental Table 1 associated with this article can be found in the online version at www.jgenetgenomics.org.

## References

**Bautista, N.S., Solis, R., Kamijima, O., and Ishii, T.** (2001). RAPD, RFLP and SSLP analyses of phylogenetic relationships between cultivated and wild species of rice. Genes Genet. Syst. **76:** 71–79.

**Becker, J., and Heun, M.** (1995). Barley microsatellites: Allele variation and mapping. Plant Mol. Biol. **27:** 835–845.

**Chang, T.T.** (1976). The origin, evolution, clutivation, dissemination and diversification of Asian and African rice. Euphytica **25:** 323–334.

**Cheng, C.Y., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H., and Ohtsubo, E.** (2003). Polyphyletic origin of cultivated rice: Based on the interspersion pattern of SINEs. Mol. Biol. Evol. **20:** 67–75.

**Dally, A.M., and Second, G.** (1990). Chloroplast DNA diversity in wild and cultivated species of rice (Genus *Oryza*, section Oryza). Cladistic-mutation and genetic-distance analysis. Theor. Appl. Genet. **80:** 209–222.

**Hampl, V., Pavlicek, A., and Flegr, J.** (2001). Construction and bootstrap analysis of DNA fingerprinting-based phylogenetic trees with the freeware program FreeTree: Application to trichomonad parasites. Int. J. Syst. Evol. Microbio. **51:** 731–735.

**Hirano, H.Y., Mochizuki, K., Umeda, M., Ohtsubo, M., and Sano, Y.** (1994). Retrotransposition of a plant SINE into the *wx* locus during evolution of rice. J. Mol. Evol. **38:** 132–137.

**Ishii T., Xu Y.B., and McCouch S.R.** (2001). Nuclear- and chloroplast-microsatellite variation in A-genome species of rice. Genome **44:** 658–666.

**Khush, G.S.** (1997). Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. **35:** 25–34.

**Li, C.B., Zhou, A.L., and Sang, T.** (2006). Rice domestication by reducing shattering. Science **311:** 1936–1939.

**Lin, H.N., Zhu, W., Silva, J.C., Gu, X., and Buell, C.R.** (2006). Intron gain and loss in segmentally duplicated genes in rice. Genome Biol. **7:** R41.

**Londo, J.P., Chiang, Y.C., Hung, K.H., Chiang, T.Y., and Schaal, B.A.** (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. Proc. Natl. Acad. Sci. USA **103:** 9578–9583.

**Lu, B.R., Zheng, K.L., Qian, H.R., and Zhuang, J.Y.** (2002). Genetic differentiation of wild relatives of rice as assessed by RFLP analysis. Theor. Appl. Genet. **106:** 101–106.

**Ma, J.X., and Bennetzen, J.L.** (2004). Rapid recent growth and divergence of rice nuclear genomes. Proc. Natl. Acad. Sci. USA **101:** 12404–12410.

**Mochizuki, K., Ohtsubo, H., Hirano, H., Sano, Y., and Ohtsubo, E.** (1993). Classification and relationships of rice strains with AA genome by identification of transposable elements at nine loci. Jpn. J. Genet. **68:** 205–217.

**Morishima, H.** (1986). Wild progenitors of cultivated rice and their population dynamics. In Rice Genetics (Manila, Philippines: International Rice Research Institute), pp. 3–14.

**Murray, M.G., and Thompson, W.E.** (1980). Rapid isolation of high-molecular-weight plant DNA. Nucleic Acids. Res. **8:** 4321–4325.

**Nei, M.** (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89:** 583–590.

**Oka, H.I.** (1974). Experimental studies on the origin of cultivated rice. Genetics **78:** 475–486.

**Oka, H.I.** (1988). Origin of Cultivated Rice. (Tokyo, Japn: Japan Scientific Societies Press).

**Oka, H.I., and Morishima, H.** (1967). Variations in the breeding systems of a wild rice, *Oryza perennis*. Evolution **21:** 249–258.

**Oka, H.I., and Morishima, H.** (1982). Phylogenetic differentiation of cultivated rice, XXIII. Potentiality of wild progenitors to evolve the *indica* and *japonica* types of rice cultivars. Euphytica **31:** 41–50.

**Rakshit, S., Rakshit, A., Matsumura, H., Takahashi, Y., Hasegawa, Y., Ito, A., Ishii, T., Miyashita, N., and Terauchi, R.** (2007). Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. Theor. Appl. Genet. **114:** 731–743.

**Rogozin, I.B., Wolf, Y.I., Sorokin, A.V., Mirkin, B.G., and Koonin, E.V.** (2003). Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. Curr. Biol. **13:** 1512–1517.

**Rose, A.B.** (2008). Intron-mediated regulation of gene expression. In Current Topics in Microbiology and Immunology, Vol. 326, Nuclear pre-mRNA Processing in Plants, A.S.N. Reddy and M. Golovkin, eds (Berlin Heidelberg: Springer-Verlag), pp. 277–290.

**Roy, S.W., and Gilbert, W.** (2005). Rates of intron loss and gain: Implications for early eukaryotic evolution. Proc. Natl. Acad. Sci. USA **102:** 5773–5778.

**Second, G.** (1982). Origin of the genic diversity of cultivated rice (*Oryza* spp.): Study of the polymorphism scored at 40 isozyme loci. Jpn. J. Genet. **57:** 25–57.

**Tang, T., Lu, J., Huang, J., He, J., McCouch, S.R., Shen, Y., Kai, Z., Purugganan, M.D., Shi, S., and Wu, C.** (2006). Genomic variation in rice: Genesis of highly polymorphic linkage blocks during domestication. PLoS Genet. **2(11):** e199.

**Ting, Y.** (1957). The origin and evolution of cultivated rice in China. Scientia Agricultura Sinica **8:** 242–260. (in Chinese with an English abstract)

**Vaughan, D.A., Morishima, H., and Kadowaki, K.** (2003). Diversity in the *Oryza* genus. Curr. Opin. Plant Biol. **6:** 139–146.

**Wang, X.K., and Sun, C.Q.** (1996). Researches on the Origin and Evolution of Chinese Cultivated Rice. (Beijing, China: China Agricultural University Press) (in Chinese).

**Wang, X.S., Zhao, X.Q., Zhu J., and Wu, W.R.** (2005). Genome-wide investigation of intron length polymorphisms and their potential as molecular markers in rice (*Oryza sativa* L.). DNA Res. **12:** 417–427.

**Wang, Z.Y., Second, G., and Tanksley, S.D.** (1992). Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of the unclear RFLPs. Theor. Appl. Genet. **83:** 565–581.

**Williams, J.G.K, Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingev, S.V.** (1990). DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids. Res. **18:** 6531–6535.

**Yang, L., Jin, G.L., Zhao, X.Q., Zheng, Y., Xu, Z.H., and Wu, W.R.** (2007). PIP: A database of potential intron polymorphism markers. Bioinformatics **23:** 2174–2177.

**Zietkiewicz, E.** (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerize chain reaction amplification. Genomics **20:** 176–183.