Rice Science, 2013, 20(2): 103−110

DOI: 10.1016/S1672-6308(13)60120-1

Comparison of Cheng's Index- and SSR Marker-based Classification of Asian Cultivated Rice

WANG Cai-hong, XU Qun, YU Ping, YUAN Xiao-ping, YU Han-yong, WANG Yi-ping, TANG Sheng-xiang, WEI Xing-hua
(*State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou 310006, China*)

Abstract: A total of 100 cultivated rice accessions, with a clear isozyme-based classification, were analyzed based on Cheng's index and simple sequence repeat (SSR) marker. The results showed that the isozyme-based classification was in high accordance with that based on Cheng's index and SSR markers. Mantel-test revealed that the Euclidean distance of Cheng's index was significantly correlated with Nei's unbiased genetic distance of SSR markers (r = 0.466, $P \leq 0.01$). According to the model-based group and cluster analysis, the Cheng's index- and SSR-based classification coincided with each other, with the goodness of fit of 82.1% and 84.7% in indica, 97.4% and 95.1% in japonica, respectively, showing higher accordance than that within subspecies. Therefore, Cheng's index could be used to classify subspecies, while SSR marker could be more efficient to analyze the subgroups within subspecies.
Key words: *Oryza sativa*; classification; Cheng's index; simple sequence repeat marker

Asian cultivated rice (*Oryza sativa* L.) holds a great proportion of variation within the species. The accurate classification of *O. sativa* is not only useful for the utilization of conventional and heterosis breeding in rice, but also for studying plant evolution. In general, the methods of rice classification involve morphological, cytological, protein and molecular markers. Using hybrid fertility as well as morphological and serological characters, Kato et al (1928) divided *O. sativa* into two main types, denoted as indica Kato and japonica Kato. Based on morphological characters including grain shape and length-width ratio, Matsuo (1952) classified rice into three types, respectively referred to as japonica, indica and javanica. However, Oka (1958) considered that the japonica and javanica should be considered as the temperate japonica and tropical japonica, respectively. According to Ding's five-level taxonomic system (Ding, 1961), rice in China was divided into indica and japonica subspecies, early-medium and late eco-groups, lowland and upland eco-types, waxy and non-waxy varieties and general cultivars. With the development of biochemical technique, intraspecific classification had also been confirmed by electrophoretically identifiable isozymes and DNA markers, etc. Glaszmann (1987) adopted fifteen polymorphic loci coding for eight isozymes and identified six varietal groups (group I to VI). Garris et al (2005) used simple sequence repeat (SSR) markers and divided Asian cultivated rice into five distinct groups, respectively as indica, aus, aromatic, tropical japonica and temperate japonica. In addition, Zhang et al (2009) revealed that indica and japonica were the two main subspecies in China. Moreover, indica showed clear differentiation among seasonal ecotypes, but not among soil-watery ecotypes; while japonica showed clear differentiations among soil-watery ecotypes, but not among seasonal ecotypes.

The traits apparent between indica and japonica, including glume hairiness, phenol reaction, interval between the 1st and 2nd nodes of panicle axis, glume color at heading, leaf pubescence and shape of grain, were identified as 'Cheng's index' (Cheng et al, 1988). The classification based on Cheng's index became easy and convenient, and was widely adopted in China. Many studies showed that the accuracy rate of Cheng's index-based classification for tropical japonica and temperate japonica could reach about 95% and 85%, respectively. According to the previous studies, the Cheng's index-based classification showed highly accordance with isozyme-based, as well as with restriction fragment length polymorphism (RFLP)-based

Received: 3 June 2012; **Accepted:** 26 September 2012
Corresponding author: WEI Xing-hua (hzxwei@gmail.com)

classification (Nakamura et al, 1992; Zhang et al, 1992).

As a new type of DNA marker, SSR, was developed in the late 1980s. Compared with isozyme and RFLP, SSR was found to be highly polymorphic, codominant, stable and repeatable, as well as easy operation. Their PCR amplifications facilitate the study of genetic diversity (Davierwala et al, 2000; Zhou et al, 2003), the construction of genome fingerprint (Zietkiewicz et al, 1994; Chakravarthi et al, 2006), the research of protection and utilization of germplasm resources (Olufowote et al, 1997; Garland et al, 1999), and so on (Zhao et al, 1992; Temnykh et al, 2001).

In this study, 100 Asian cultivated rice accessions, with a clear isozyme-based classification, were used. The object was to compare the Cheng's index- and SSR marker-based classification, and reveal the relationship between Cheng's index and SSR markers.

MATERIALS AND METHODS

Rice materials

According to the isozyme-based classification of Glaszmann (1987), a total of 100 rice accessions were chosen from the Genetic Resources Center, International Rice Research Institute, Manila, Philippines. Accession name, origin and classification are shown in Supplemental Table 1.

Investigation for six characters of Cheng's index

All accessions were grown in fields in Lingshui, Hainan Province, China. Six rows with six plants were planted for each accession, one per bunch and arranged orderly (for discerning similar characters). After random sampling by disgorging the plants in the fringe area, six characters of 100 accessions were investigated as follows and converted into the score value according to the standards of Cheng's index (Cheng , 1993).

1) Glume hairiness: take five grains randomly, and observe the morphological characters of glume hairness;

2) Phenol reaction: take five grains randomly, and soak them into 2% phenol solution in the petri dish for about 3–5 d, then observe the glume color;

3) Interval between the 1st and 2nd nodes of panicle axis: measure five panicles randomly between the 1st and 2nd nodes;

4) Glume color at heading: take five spikes randomly, and identify the glume color at the heading stage;

5) Leaf pubescence: take five leaves randomly, and judge how many leaf pubescences at the heading stage;

6) Shape of grain: take five grains randomly, measure their lengths and widths, then calculate the length-width ratio.

DNA extraction

For each accession, five seeds were germinated in petri plates at ambient temperature under natural light. DNA was extracted from 1–2 cm from the top of tender shoot following a modified sodium dodecyl sulfate (SDS) mini-extraction protocol developed by Zheng et al (1995), quantified using a spectrophotometer, and checked by 1.5% agarose gel electrophoresis, then dissolved in 0.1 × TE solution and stored at -20 ºC.

Fluorescent SSR markers

One hundred and nine SSR markers (Supplemental Table 2) were randomly selected (7–10 markers on each of the 12 chromosomes). The primer sequences were synthesized by ABI (Applied Biosystems, Foster City, USA). For each SSR, the forward primer was labeled with a fluorescent label.

PCR amplification

DNA amplification was carried out in a 2720 thermal cycler (Applied Biosystems, Foster City, USA) in a 10 μL reaction mixture. Each reaction contained 10 × buffer 1.0 μL, 2 mmol/L dNTPs 1.0 μL, 25 mmol/L $MgCl_2$ 1.0 μL, 0.6 μL each of the forward and reverse primers (10 μmol/L), 5 U/μL *Taq* polymerase 0.1 μL, and 20 ng of template DNA. The thermal cycling profile was as follows: 94 ºC for 2 min; 35 cycles of 94 ºC for 45 s, 55 ºC (50 ºC for RM6091, 61 ºC for RM161 and RM185, 67 ºC for RM119, RM132, RM169 and RM176, respectively) for 45 s, and 72 ºC for 1 min; and 72 ºC for 8 min.

Products test

The PCR products (2 μL) were diluted in 8 μL of ultra pure water, and purified in the solution mixed with ethnol and sodium acetate for 30 min; then scoured in 100% alcohol for 15 min. A diluted DNA was dried and mixed with highly deionized-formamide (Applied Biosystems, Foster City, USA), then submitted for fragment analysis by capillary electrophoresis on an Applied Biosystems 3130xl DNA analyzer (Applied Biosystems, Foster City, USA).

Data analysis

DNA fragment size analysis was performed using GeneMapper software, followed by manual allele

binning. STUCTURE was used to infer clusters of similar SSR genotypes. STATISTICA 7.0 was performed on the six morphological traits of Cheng's index according to Euclidean distance. PowerMarker 3.25 was applied to construct the Unweighted Pair-Group Method with Arithmetic means (UPGMA) dendrogram on SSR genotypes according to Nei's unbiased genetic distance. The correlation coefficient between Cheng's index- and SSR marker-based genetic distance matrices was calculated in Mantel-test using ARLEQUIN 3.0.

RESULTS

Comparison of Cheng's index- and isozyme-based classification

Accessions in this study were divided into six groups by Glaszmann (1987). As shown in Supplemental Table 1, accessions in groups I, II and III were classified as C-H or C-H′, while those in groups IV, V and VI were as C-K′ or C-K, with a goodness of fit reaching 74.1% and 76.1%, respectively (Table 1). Among these groups, accessions belonged to group I and group VI-A had the highest goodness of fit with C-H or C-H′ and C-K′ or C-K (87.5% and 80.0%, respectively), while group II had the lowest goodness of fit with C-H or C-H′ as 59.1% (Table 1).

Comparison of SSR marker- and isozyme-based classification

The model-based simulation of population structure using SSR markers showed that 100 accessions were divided into six groups (Table 1). Compared with isozyme- based classification, accessions in groups I′ and II′ corresponded to groups I, II and III, while those in groups III′, IV′, V′ and VI′ corresponded to groups IV, V and VI, with a goodness of fit reaching 90.7% and 80.4%, respectively (Table 1). Among these groups, I′ corresponded to I, II′ to II and III, III′ to IV, IV′ to V, V′ to VI-A, and VI′ to VI-B, with the goodness of fit reaching 100%, 63.3%, 72.7%, 41.7%, 73.3% and 37.5%, respectively (Table 1). However, JC73-4 (group II), JC101 (group II), KUMRI BORO (group III) and KALIJIRA (group IV) were classified into group IV′ according to SSR marker-based classification. Furthermore, Bhadoia 233, Phudugey, Basmati 1, IR8-246/NC1626CN2-5-9-4, RAYADA, Basmati, Som Cau 70A, Yelaik Meedon, Baghlani Nangarhar and KU115 contained more than one component, and were defined as admixture accessions (Supplemental Table 1).

Comparison of Cheng's index- and SSR marker-based classification

Mantel-test showed that the correlation coefficient between Cheng's index- and SSR marker-based genetic distance matrices was 0.466 ($P \leq 0.01$), which manifested that a certain number of morphological characters with strong specificity between indica and japonica and could be used to identify these two subspecies successfully.

Classification based model

When $K = 2$, 100 accessions were divided into two groups, S-H or S-H′ and S-K′ or S-K, which were respectively corresponded to C-H or C-H′ and C-K′ or

Table 1. Goodness of fit of classification based on isozyme, Cheng's index and SSR marker. %

Classification		Isozyme-based								
		I, II, III	IV, V, VI	I (24)	II (22)	III (8)	IV (11)	V (12)	VI (23)	
									VI-A (15)	VI-B (8)
Cheng's index-based	C-H (22) or C-H′ (29)	74.1		87.5	59.1	75.0				
	C-K′ (22) or C-K (27)		76.1				72.7	75.0	80.0	62.5
SSR marker-based	I′ (30)	90.7		100						
	II′ (22)				63.3					
	III′ (8)		80.4				72.7			
	IV′ (9)							41.7		
	V′ (12)								73.3	
	VI′ (8)									37.5

Numbers in the parenthesis are the numbers of accessions. C-H, Hsien based on Cheng's index; C-H′, Hsien-cline based on Cheng's index; C-K, Keng based on Cheng's index; C-K′, Keng-cline based on Cheng's index.

Classification based on isozyme (Glaszmann, 1987): I corresponds to Hsien ecotype; II corresponds to Aus ecotype; III consists of short cycle, photoperiod insensitive and deep water rice; IV corresponds to the Rayada rices; V is considered as Aromatic ecotype; VI is dominant in temperate areas (VI-A) and in high elevation areas (VI-B). Classification based on SSR markers (Garris et al, 2005): I′ corresponds to indica group; II′ corresponds to aus group; IV′ corresponds to aromatic group; V′ corresponds to temperate japonica; VI′ corresponds to tropical japonica; III′ corresponds to rayada group compared with isozyme-based classification (Glaszmann, 1987).

Table 2. Comparison of Cheng's index- and SSR marker-based classification according to model.

Parameter	Group					
	K = 2		*K* = 4			
Cheng's index	C-H or C-H′	C-K′ or C-K	C-H	C-H′	C-K′	C-K
Number of accessions	51	49	22	29	22	27
SSR marker	S-H or S-H′	S-K′ or S-K	S-H	S-H′	S-K′	S-K
Number of accessions	58	39	36	22	21	18
Number of common accessions	48	38	18	11	8	14
Goodness of fit (%)	82.1	97.4	50.0	50.0	38.1	76.5

C-H, Hsien based on Cheng's index; C-H′, Hsien-cline based on Cheng's index; C-K, Keng based on Cheng's index; C-K′, Keng-cline based on Cheng's index.

Table 3. Comparison of Cheng's index- and SSR markers-based classification according to genetic distance.

Classification	No. of accessions		No. of common accessions [a]	Goodness of fit (%)
	Cheng's index-based	SSR marker-based		
Group I	52	59	50	84.7
Subgroup 1	29	37	21	56.8
Subgroup 2	23	22	10	45.5
Group II	48	41	39	95.1
Subgroup 3	15	24	7	29.2
Subgroup 4	33	17	12	70.6

When the Euclidean distance based on Cheng's index was set at 6.52, or Nei's unbiased genetic distance based on SSR marker was set at 0.39, the accessions were divided into groups I and group II, referring to indica and japonica, respectively. When the Euclidean distance was set at 5.61 or Nei's unbiased genetic distance based on SSR marker was set at 0.38, the accessions were divided into four subgroups, among which subgroups 1 and 2 responded to group I and subgroups 3 and 4 responded to group II.

[a] The accession classified into the same group by both Cheng's index- and SSR marker-based classification is regarded as common accession.

C-K (goodness of fit = 82.1% and 97.4%) (Table 2). When *K* = 4, referring to S-H, S-H′, S-K′ and S-K, the goodness of fit between S-K and C-K was the highest (76.5%), followed by the goodness of fit between S-H and C-H, and between S-H′ and C-H′ (50.0%), while the goodness of fit between S-K′ and C-H′ was the lowest (38.1%). In addition, Bhadoia 233, Phudugey, IR8-246/NC1626CN2-5-9-4, Basmati, Yelaik Meedon and KU115 were defined as admixture accessions (Supplemental Table 1).

Classification based on Euclidean distance and genetic distance

When the Euclidean distance, based on Cheng's index, was set at 6.52, 100 accessions were divided into two groups, indica and japonica. In addition, when the Nei's unbiased genetic distance based on SSR markers was set at 0.39, there were also two groups. By comparision of the results from classification, it could be found that these two subspecies were accordant at the goodness of fit of 84.7% and 95.1%, respectively (Table 3).

It had been shown that each subspecies could be divided into two subgroups when the Euclidean distance was 5.61 and the Nei's unbiased genetic distance was 0.38. The goodness of fit among these four subgroups was different from each other, respectively 56.8%, 45.5%, 29.2% and 70.6% (Table 3). In comparison with isozyme-based classification, Cheng's index could not evidently classify the subgroups within the subspecies (Fig. 1) which was accurately and clearly reflected by SSR markers (Fig. 2).

DISCUSSION

Zhou et al (1988) calculated the principal coordinate analysis and cluster dendrogram with 197 accessions from Glaszmann, and demonstrated that Cheng's index-based classification was in high agreement with isozyme-based results. In our study, Cheng's index- and isozyme-based classification was also in a goodness of fit reaching more than 70%. Moreover, accessions in group I had the highest goodness of fit with C-H or C-H′ (87.5%). Our studies also showed that isozyme- and SSR marker-based classification was in good accordance.

In addition, the correspondence existed between morphological traits and SSR markers (Wu et al, 2001; Yu et al, 2004). In this study, the mantel-test showed that the Euclidean distance matrix based on Cheng's index was highly and significantly correlated with the Nei's unbiased genetic distance matrix based on SSR markers, which reflected that a certain number of morphological characters with strong specificity could be used to identify indica and japonica. Ma et al (2010)

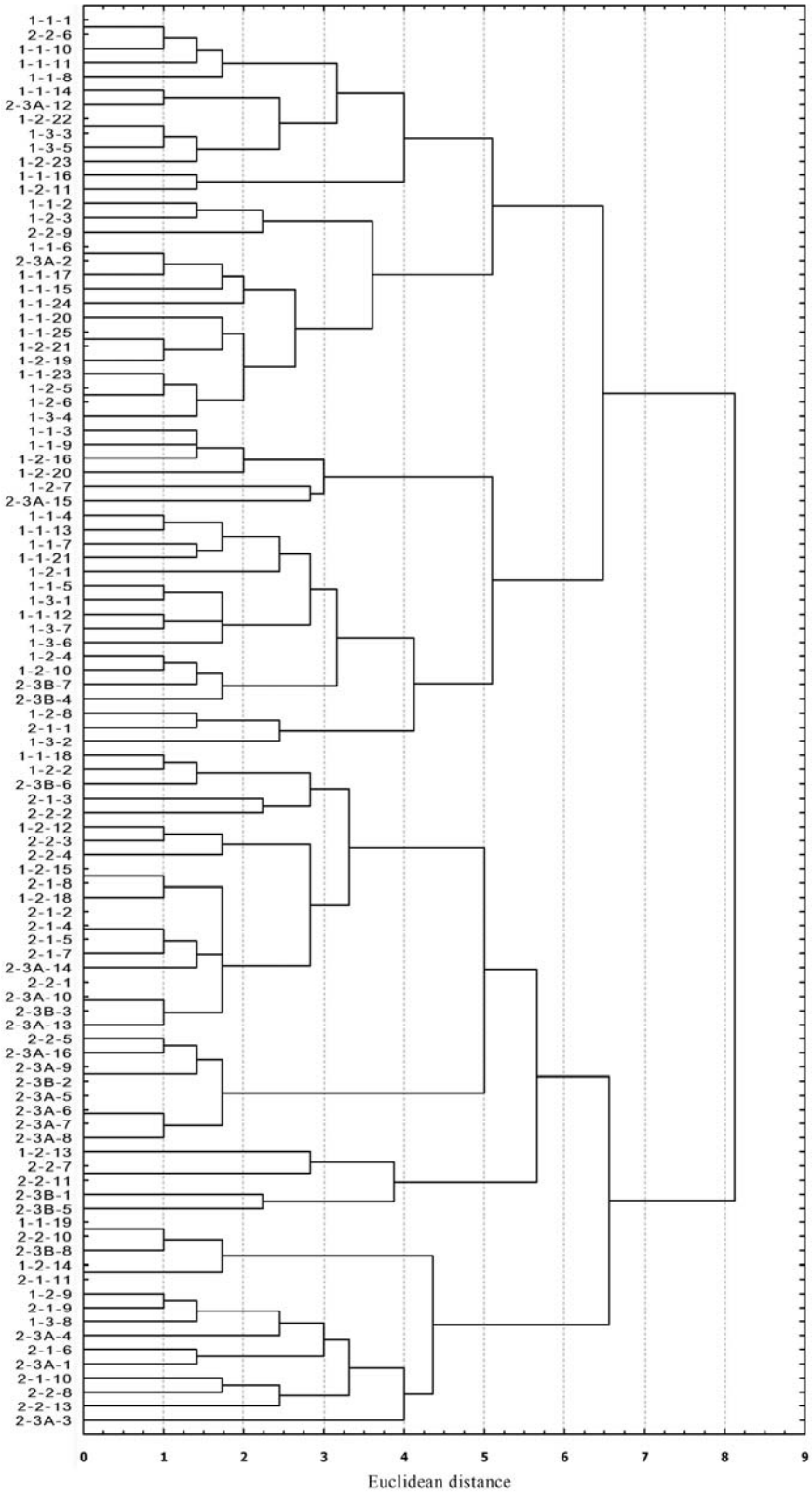

Fig. 1. Unweighted Pair-Group Method with Arithmetic (UPGMA) dendrogram of Cheng's index-based classification according to Euclidean distance.

When the value of Euclidean distance was 6.52, the accessions were divided into two gruops; when the value was 5.61, the accessions were divided into four subgroups.

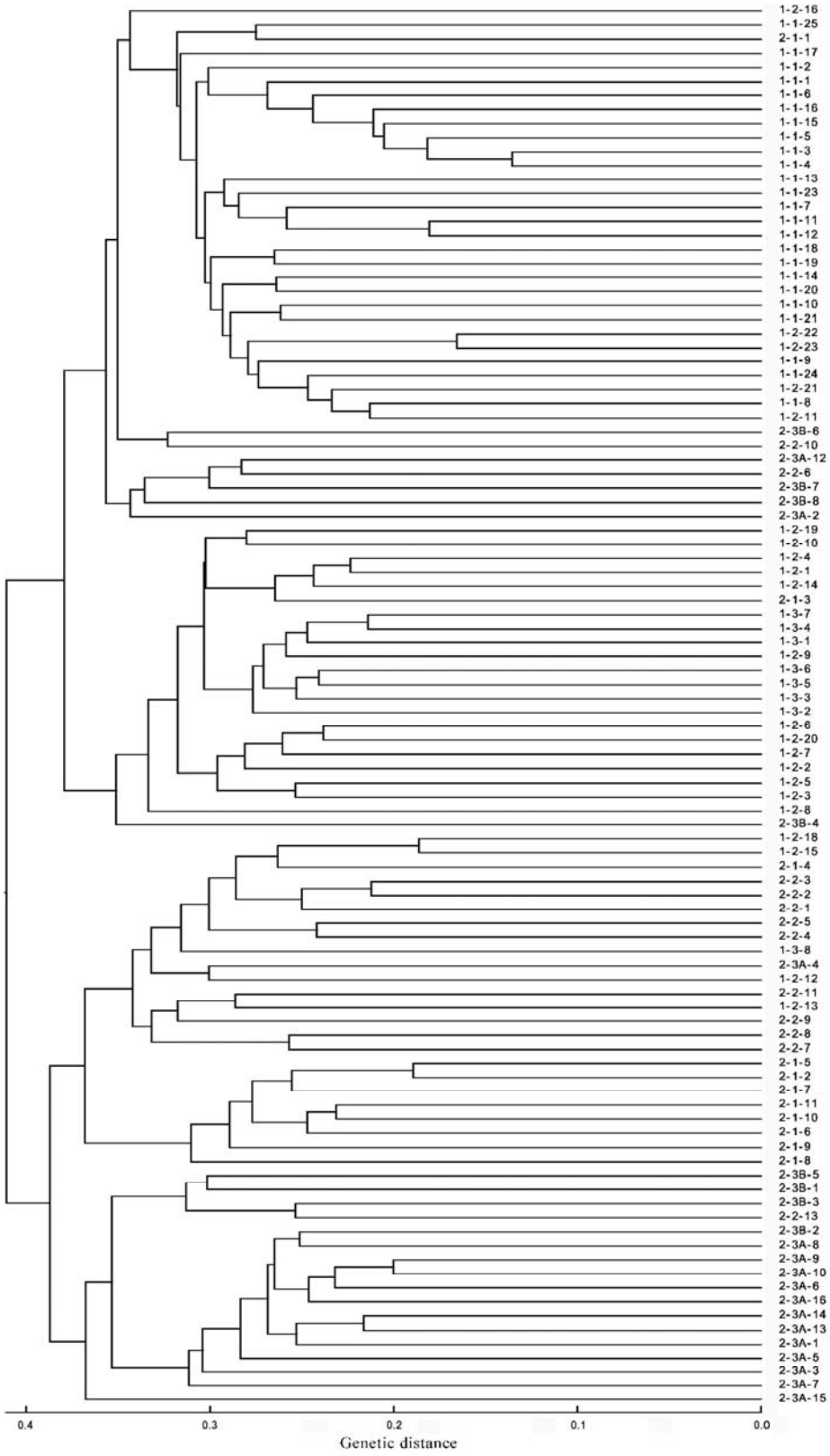

Fig. 2. Unweighted Pair-Group Method with Arithmetic (UPGMA) dendrogram of SSR marker-based according to Nei's unbiased genetic distance.

When the value of Nei's unbiased genetic distance was set at 0.39, the accessions were divided into two groups; when the value was 0.38, the accessions were divided into four subgroups.

found that Cheng's index had a little difference with SSR markers when they were used to analyze the genetic structure of Guizhou landrace. Compared with SSR marker-based classification, isozyme-based classification had a lower goodness of fit with that based on Cheng's index, and it may be due to isozyme had fewer loci and less genetic information. The results of structure analysis based on model and genetic distance both showed that Cheng's index- and SSR marker-based classification had a higher goodness of fit among the subspecies than within subspecies. Therefore, morphological characters of Cheng's index had strong specificity between indica and japonica, while SSR markers were more efficient to analyze the subgroups within subspecies.

The contribution rates of morphological characters used in Cheng's index-based classification are not equal and often make a certain deviation. Zhou et al (1988) found that the contribution rate of glume hairiness was the highest (81.5%), followed by glume color at heading (80.6%), shape of grain and interval between the 1st and 2nd nodes of panicle axis (54.4%). Zhang et al (1998) reported the deviation of 5%–10% for classifying rice varieties between the Cheng's index-based and isozyme-based classifications, and considered that glume hairiness had the highest goodness of fit (93.0%), followed by glume color at heading (83.0%), phenol reaction (81.0%), leaf hairiness (77.0%), interval between the 1st and 2nd nodes of panicle axis (69.0%) and shape of grain (46.0%) in turn. In our study, Taducan was classified as indica by isozyme- and SSR marker-based classification, but as japonica by Cheng's index-based classification. It might be related with the high contribution rate of glume hairiness, glume color at heading, phenol reaction and leaf hairiness. In addition, JATRA MOTUK was classified as japonica by isozyme- and SSR marker-based analysis, but classified as indica by Cheng's index-based analysis. It might be explained with high contribution rates of glume color at heading, phenol reaction and interval between the 1st and 2nd nodes of panicle axis. Therefore, we should pay attention to distinguish the deviation due to the different contribution rates of morphological characters in Cheng's index-based classification to identify indica and japonica.

ACKNOWLEDGEMENTS

The authors are grateful to the International Rice Research Institute, the Philippines for providing the accessions used in this study. This work was supported by the Crop Genetic Resources Protection Project of Ministry of Agriculture, China, and the Basic Research Budget of China National Rice Research Institute (Grant No. 2009RG001-3).

SUPPLEMENTAL DATA

The following materials are available in the online version of this article at http://www.sciencedirect.com/science/journal/16726308; http://www.ricescience. org.

Supplemental Table 1. Accession name, origin and classification of accessions used in this study.

Supplemental Table 2. Identities of 109 SSR loci used for analysis, including chromosomal location and primer sequence.

REFERENCES

Chakravarthi B K, Naravaneni R. 2006. SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L). *Afr J Biotechn*, **5**(9): 684–688.

Cheng K S. 1988. A statistical evaluation of the classification of rice cultivars into *hsien* and *keng* subspecies. *Rice Genet News*, **4**: 46–48.

Cheng K S. 1993. Discrimination of Indica and Japonica Subspecies in Asian Cultivated Rice. Kunming: Yunnan Scientific & Technical Publishers: 6–7. (in Chinese)

Davierwala A P, Chowdari K V, Kumar S, Reddy A P K, Ranjekar P K, Gupta V S. 2000. Use of three different marker systems to estimate genetic diversity of Indian elite rice varieties. *Genetica*, **108**: 269–284.

Ding Y. 1961. Rice Cultivation in China. Beijing: Agriculture Publisher: 83–184. (in Chinese)

Garland S H, Lewin L, Abedinia M, Henry R, Blakeney A. 1999. The use of microsatellite polymorphisms for the identification of Australian breeding lines of rice (*Oryza sativa* L.). *Euphytica*, **108**: 53–63.

Garris A J, Tai T H, Coburn J, Kresovich S, McCouch S. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics*, **169**: 1631–1638.

Glaszmann J C. 1987. Isozymes and classification of Asian rice varieties. *Theor Appl Genet*, **74**: 21–30.

Kato S, Kosaka H, Hara S. 1928. On the affinity of rice varieties as shown by the fertility of hybrid plants. *Bull Sci Fac Agric Kyushu Univ*, **3**: 132–147.

Ma L, Yu X Q, Zhao F S. 2010. SSR-based analysis on genetic diversity of landraces from Guizhou Province, China. *Chin J Rice Sci*, **24**(3): 237–243. (in Chinese with English abstract)

Matsuo T. 1952. Genecological studies on cultivated rice. *Bull Natl Inst Agric Sci*, **D3**: 1–111. (in Japanese)

Nakamura M, Yoshimura A, Iwata N. 1992. Phylogenetic study of cultivated rice and its wild relatives by RFLP. *Rice Genet News*,

19: 132–134.

Oka H I. 1958. Intervarietal variation and classification of cultivated rice. *Ind J Genet Plant Breed*, **18**: 79–89.

Olufowote J O, Xu Y B, Chen X L, Goto M, McCouch S R, Park W D, Beachell H M, Dilday R H. 1997. Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. *Genome*, **40**: 370–378.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Res*, **11**: 1441–1452.

Wu Y T, Zhang T Z, Yin J M. 2001. Genetic diversity detected by DNA markers and phenotypes in upland cotton. *Acta Genet Sin*, **28**(11): 1040–1050. (in Chinese with English abstract)

Yu H Y, Wei X H, Wang Y P, Tang S X. 2004. Study on genetic variation of rice varieties derived from Aizizhan by using morphological traits, allzoymes and simple sequence repeat (SSR) markers. *Chin J Rice Sci*, **18**(6): 477–482. (in Chinese with English abstract)

Zhang Q F, Maroof M A S, Lu T Y, Shen B Z. 1992. Genetic diversity and differentiation of indica and japonica rice detected by RFLP analysis. *Theor Appl Genet*, **83**: 495–499.

Zhang Y Z, Xu N. 1998. A discussion on the method and system of rice indica-japonica classification. *Southwest Chin J Agric Sci*, **11**(3): 88–93. (in Chinese with English abstract)

Zhang D L, Zhang H L, Wang M X, Sun J L, Qi Y W, Wang F M, Wei X H, Han L Z, Wang X K, Li Z C. 2009. Genetic structure and differentiation of *Oryza sativa* L. in China revealed by microsatellites. *Theor Appl Genet*, **119**(6): 1105–1117.

Zhao X P, Kochert G. 1992. Characterization and genetic mapping of a short, highly repeated, interspersed DNA sequence from rice (*Oryza sativa* L). *Theor Appl Genet*, **231**: 353–359.

Zheng K L, Subudi P K, Domingo J, Magpantay G, Huang N. 1995. Rapid DNA isolation for marker assisted selection in rice breeding. *Rice Genet News*, **12**: 255–258.

Zhou H F, Xie Z W, Ge S. 2003. Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (*Oryza rufipogon* Griff) in China. *Theor Appl Genet*, **107**: 332–339.

Zhou H, Glasamann J C, Cheng K S, Shi X Q. 1988. A comparison of methods in classification of cultivated rice. *Chin J Rice Sci*, **2**(1): 1–7. (in Chinese with English abstract)

Zietkiewicz B, Rafalski A, Labuda D. 1994. Genome fingerprint by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. *Genomics*, **20**: 176–183.