NOTES

Rapid and reliable identification of rice genomes by RFLP analysis of PCR-amplified *Adh* genes

Song Ge, Tao Sang, Bao-rong Lu, and De-yuan Hong

Abstract: The rice genus (*Oryza* L.) consists of 24 species with 10 recognized genome types. With the realization of many useful genes in species of wild rice, continuous efforts have been made to understand their genomic composition and relationships. However, the identification of rice genomes has often been difficult owing to complex morphological variation and formation of allotetraploids. Here we propose a rapid and reliable method for identifying rice genomes based on the restriction sites of PCR-amplified *Adh* genes. The experimental procedure was as follows: (*i*) amplify a portion of *Adh1* and *Adh2* genes with the locus-specific PCR primers; (*ii*) digest PCR products with restriction enzymes that distinguish different genomes; and (*iii*) run the digested products on 1.4% agarose gel, and photograph. Using various combinations of restriction digestion of the two *Adh* genes, all of the rice genomes can be identified.

Key words: *Adh* gene, genome, identification, *Oryza* L., PCR–RFLP.

Résumé : Le riz appartient au genre *Oryza* L., lequel comprend 24 espèces et 10 types génomiques reconnus. Étant donné l'existence de nombreux gènes utiles au sein d'espèces sauvages de riz, des travaux ont été réalisés d'une façon soutenue en vue d'acquérir une meilleure compréhension de la composition génomique de ces divers riz et des relations entre ceux-ci. Cependant l'identification des génomes du riz s'avère souvent difficile en raison d'une variation morphologique complexe et de la formation d'allotétraploïdes. Les auteurs proposent ici une méthode rapide et fiable pour identifier les génomes du riz en s'appuyant sur l'analyse des sites de restriction présents dans des amplicons issus des gènes *Adh*. La procédure expérimentale est la suivante : (*i*) amplifier un segment des gènes *Adh1* et *Adh2* à l'aide d'amorces spécifiques de chaque locus; (*ii*) digérer les amplicons avec des enzymes de restriction qui permettent de distinguer les différents génomes; (*iii*) séparer les produits de digestion sur un gel d'agarose 1,4 % et photographier. En utilisant diverses combinaisons d'enzymes de restriction sur les deux gènes *Adh*, tous les génomes du riz peuvent être identifiés.

Mots clés : gène *Adh*, génome, identification, *Oryza* L., PCR–RFLP.

[Traduit par la Rédaction]

Introduction

The genus *Oryza* L. consists of 2 cultivated and approximately 22 wild species. The Asian cultivated rice (*O. sativa* L.) is an economically important world crop that provides staple food for more than one half of the global population (Lu et al. 2000). With the realization of the potential agricultural value of the wild rice species, continuous efforts have been made to understand the genomic composition and evolutionary relationships within the genus *Oryza* (Nayar 1973; Jena and Kochert 1991; Cordesse et al. 1992; Sarkar and Raina 1992; Wang et al. 1992; Aggarwal et al. 1997; McIntyre and Winberg 1998; Ge et al. 1999). Based on classical genomic analyses using cytogenetic methods and recent techniques of DNA hybridization and sequencing, 10 genome types (A, B, BC, C, CD, E, F, G, HJ, and HK) were recognized for the *Oryza* species (Vaughan 1994; Aggarwal et al. 1997; Khush 1997; Ge et al. 1999; Lu et al. 2000).

Genomic identification of a given rice sample is very important for its correct classification, particularly for tetraploid species. However, genome identification by classical cytological methods is complicated and time consuming, and that based on morphology is very difficult owing largely to the lack of reliable qualitative characters. Various molecular approaches have been proposed for the determination of rice genomes and species, such as tandemly repeated se-

Received April 6, 2001. Accepted June 29, 2001. Published on the NRC Research Press Web site at http://genome.nrc.ca on November 7, 2001.

Corresponding Editor: K.J. Kasha.

S. Ge[1] **and D. Hong.** Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China.
T. Sang. Department of Botany and Plant Pathology, Michigan State University, East Lansing, MI 48824, U.S.A.
B. Lu. Institute of Biodiversity Science, School of Life Sciences, Fudan University, Shanghai 200433, China.

[1]Corresponding author (e-mail: gesong@ns.ibcas.ac.cn or song_ge@hotmail.com).

 DOI: 10.1139/gen-44-6-1136

quences (Zhao et al. 1989; De Kochko et al. 1991; Reddy et al. 1993), DNA spacer fragments (Cordesse et al. 1992), and seed-protein profile (Sarkar and Raina 1992). Recently, total genomic DNA hybridization and dot hybridization have also been used to identify rice genomes (Aggarwal et al. 1997; McIntyre and Winberg 1998). These methods either failed to identify the genomes unambiguously or involved relatively complex experimental procedures. Here we describe a quick and reliable method to identify rice genomes, i.e., the restriction fragment length polymorphism (RFLP) analysis of PCR-amplified *Adh* genes.

RFLP analysis of PCR-amplified chloroplast DNA (cpDNA) regions has been used to identify seed parents and estimate phylogenetic relationships in many plant groups (Arnold et al. 1991; Liston 1992; Badenes and Parfitt 1995; Isshiki et al. 1998). However, the RFLP analysis of PCR-amplified nuclear DNA regions has not been widely utilized. According to the phylogenetic study of the rice genus (Ge et al. 1999), the amount of sequence variation in the *Adh* genes seems to be sufficient to allow the development of RFLP markers for the identification of the rice genomes. The primary goal of this study, therefore, is to develop an approach of RFLP analysis of low-copy nuclear genes to identify the rice genomes quickly and reliably. This would greatly facilitate the identification of collected wild rice germplasm, and potentially be applied to effective germplasm collection and identification of other crops.

Materials and methods

Plant materials

All of the *Oryza* species representing the 10 rice genomes were sampled for this study. These species are listed in Table 1 with their respective accession numbers, genome constitutions, and origins. Plants were grown from seeds provided by the International Rice Research Institute (IRRI) in the Philippines except for *O. schlechteri*, which was provided in the form of dried leaves by IRRI.

DNA extraction and amplification of specific *Adh* regions

DNA was isolated from fresh or silica-gel-dried leaf material using the method of Doyle and Doyle (1987). Regions including the *Adh1* and *Adh2* genes were amplified using a universal forward primer AdhF1 (5′-CACACCGACGTCTACTTCTG-3′), and specific reverse primers, Adh1bR (5′-TCAGCAAGTACCTAAATTATC-3′) for *Adh1* and Adh2RR (5′-CCACCGTTGGTCATCTCAAT-3′) for *Adh2*, respectively. The PCRs were carried out on a PTC-200 system (MJ Research Inc., Watertown, Mass.), and included the following cycles: (1) 70°C for 4 min; (2–4) 94°C for 1 min, 56°C for 30 s, and 72°C for 1.5 min; (5–34) 94°C for 20 s, 55°C for 20 s, and 72°C for 1.5 min; (35) 72°C for 10 min. Reactions were carried out in a 25-μL mixture containing 10 mM Tris-HCl (pH 8.2); 1.5 mM $MgCl_2$; 0.2 mM each of dATP, dCTP, dGTP, and dTTP; 0.2 μM of random primer; 20 ng of genomic DNA; and 1.0 U of Taq polymerase. The PCR products were examined with 1% agarose gel in TBE buffer.

Selection of restriction enzymes

In our previous phylogenetic study, we have sequenced a portion of the transcribed regions of two nuclear genes (*Adh1* and *Adh2*) and one chloroplast gene (*matK*) for all of the rice species (Ge et al. 1999). The sequences of the two *Adh* genes for all of the 23 *Oryza* species were deposited in the GenBank database under accession numbers AF148568–AF148635. To establish a system to identify all of the rice genomes, we have surveyed the restriction sites of the two *Adh* genes extensively using DNA Strider 1.2 (Marck 1988). The enzymes that had restriction sites conserved among species within a given genome and variable among different genomes were chosen for RFLP analyses.

Two to 5 μL of PCR products were digested in a 10-μL reaction containing 2 U of restriction enzymes. Digested PCR products were electrophoresed on 1.4% TBE agarose gels. The gels were stained with ethidium bromide, visualized, and photographed under UV light.

Results

Amplification and characteristics of the target regions of two *Adh* genes

Using the specific primers, we amplified a single fragment for all the species, ranging from 1.8 to 2.0 kb in length for *Adh1* and from 1.8 to 2.4 kb in length for *Adh2*. This region spans from intron 2 to 7, and contains five exons and six introns (Fig. 1). In comparison to *Adh1*, *Adh2* shows length variation among rice genomes. Longer length in the *Adh2* region of the F genome (Fig. 2*c*) is due to a large insertion in intron 3.

Strategy for determining genome-specific sites

In our phylogenetic analysis, we have verified nine known genomes and designated a new genome type (HK) (Ge et al. 1999). The allotetraploid species had two distinct types of sequences corresponding to their respective diploid parents. For instance, two types of sequences of the BBCC species form clades with diploid species of the B and C genomes, respectively. Consequently, for both *Adh1* and *Adh2* loci, all of the sequences of *Oryza* species formed monophyletic groups according to their genome types.

Using DNA Strider, we surveyed 170 enzymes for the restriction sites of the two *Adh* genes of *Oryza* species. Of them, 40 enzymes (25 for *Adh1* and 24 for *Adh2*) showed one to three restriction sites which were conserved across species with the same genome types but variable among different genome types. Finally, 19 enzymes have at least one restriction site that was unique to between one and three genomes, and thus were chosen for subsequent RFLP analysis.

Identification of rice genomes

A total of 22 combinations of the two *Adh* genes and the 19 enzymes can be used to identify rice genomes (Table 2). The patterns of restriction fragments can be categorized into two types. The first type involved the combinations that have no cutting site for the target genome(s), but at least one cutting site for the other genomes (Nos. 1, 12, 19, and 21 in Table 2). For example, enzyme *Tth*111I has no cutting site in *Adh1* in the A genome but one or two cutting sites in the remaining genomes, resulting in only one band for the A genome species and two or three bands for the others (No. 1 in Table 2).

The second type included most of the combinations that have at least one cutting site in the target genome but no site in the remaining genomes. Of those, 10 combinations (Nos. 4–6, 10, 13, 14, 16, 17, 20, 22) were specific to a single genome, 6 were specific to two genomes (Nos. 2, 3, 7, 8,

Table 1. List of species studied, their genomic constitution, accession numbers, and origins.

Species	Genome	Accession No.	Locality
O. barthii	AA	104140	Cameroon
O. glaberrima	AA	104042	Chad
O. glumaepatula	AA	100968	Suriname
O. longistaminata	AA	104977	Kenya
O. meridionalis	AA	103317	Australia
O. nivara	AA	106148	Laos
O. rufipogon	AA	105942	Thailand
O. sativa	AA	Au73030	China
O. punctata	BB	104017	Cameroon
O. officinalis	CC	105085	Philippines
O. rhizomatis	CC	105448	Sri Lanka
O. eichingeri	CC	101422	Uganda
O. punctata	BBCC	104059	Nigeria
O. minuta	BBCC	101082	Philippines
O. alta	CCDD	105143	Guyana
O. grandiglumis	CCDD	105669	Brazil
O. latifolia	CCDD	105141	Costa Rica
O. australiensis	EE	105263	Australia
O. brachyantha	FF	105151	Sierra Leone
O. ridleyi	HHJJ	100877	Malaysia
O. longiglumis	HHJJ	105148	Indonesia
O. schlechteri	HHKK	82047	Papua New Guinea
O. granulata	GG	106469	Vietnam
O. meyeriana	GG	104987	Malaysia

11, 15), and 1 was specific to three genomes (No. 9). In most cases where the cutting sites were shared by two or three genomes, there were basically different restriction profiles for different genomes. For example, although there was one cutting site in both the A and the F genomes in the *Adh2* + *Sac*II combination (No. 2), the two genomes can be easily distinguished from each other by their restriction patterns. The two bands of the A genome species are 0.82 and 0.89 kb, whereas the F genome has two bands of 0.83 and 1.53 kb (Fig. 2*a*).

According to the combinations listed in Table 2, we could identify each of 10 genomes unambiguously by one of the combinations of *Adh* genes and restriction enzymes. To make the identification more practical, we have chosen six enzymes that are less expensive and relatively easier to obtain. The cutting sites of these enzymes and their approximate positions in the *Adh* genes are illustrated in Fig. 1. By digesting either of the *Adh* genes with these enzymes, we obtained six combinations (bold faced in Table 2) by which all of the rice genomes can be easily identified. As shown in Fig. 1, the A and F genomes could be identified by the *Adh2* + *Sac*II combination, with both of them having two bands with different profiles (Fig. 2*a*). Combinations *Adh1* + *Sac*II, *Adh2* + *Eco*NI, and *Adh1* + *Ban*II identified the B, C, and G genomes, respectively, (Fig. 2*b*–2*d*). The D(E) and HJ genomes could be identified with the combination *Adh1* + *Afl*II by their different restriction profiles (Fig. 2*e*). The *Adh2* + *Bst*EII combination showed four bands for the HK genome but one or two bands for the remaining genomes (Fig. 2*f*). This identification system worked well for all the species we tested.

Identification of the genome constitution of tetraploid species

Of 10 genomes in the genus *Oryza*, 4 originated from tetraploid species. An allotetraploid genome should display restriction patterns of both the diploid genomes that it contains. Using the combinations in Table 2, tetraploid genomes can be distinguished from the closely related diploid genomes, such as B vs. BC (Fig. 2*b*) and C vs. BC and CD (Fig. 2*c*).

Discussion

Although the potential use of wild relatives of crop species in plant breeding has been widely appreciated (Tanksley and McCouch 1997), the management of the germplasm collections has been difficult because of their sizes and the problems related to the accurate identification of collected materials. Therefore, accurate determination of rice germplasm is essential for effective conservation and efficient utilization of genetic resources in the rice gene pool.

For various reasons, the initial designation or field determination of rice germplasm accessions stored in gene banks may not always be reliable (Transue et al. 1994; Virk et al. 1995). For the genus *Oryza*, there is a large germplasm collection (about 100 000 accessions stored at the International Rice Geneback (IRRI) alone). However, the identification of rice germplasm is problematic (Virk et al. 1995). In the comprehensive study by Wang et al. (1992) on the phylogenetic relationships of 21 *Oryza* species using nuclear RFLPs, they found that 12 out of 93 accessions tested were not correctly identified, including species with the A, C, BC, and CD

Fig. 1. Maps of restriction sites of *Adh* genes for the five enzymes. The genome(s) with a given restriction site are in parentheses. The minus sign before the genome letter indicates a genome without a given restriction site that occurs in other genomes. The cutting sites for *Adh1* are above the gene, and the sites for *Adh2* are below the gene. Arrows indicate locations and directions of the PCR primers. AdhF1 is a universal primer applied to both *Adh1* and *Adh2*. Adh1bR and Adh2RR are specific primers for *Adh1* and *Adh2*, respectively. Note that some genomes may share identical restriction sites but have different restriction profiles because of length variation (see Table 2).

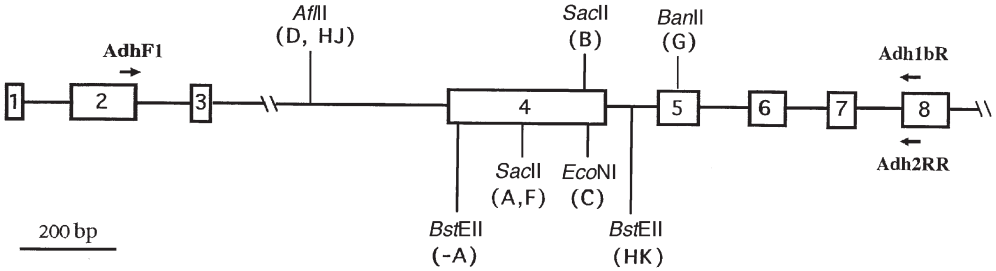
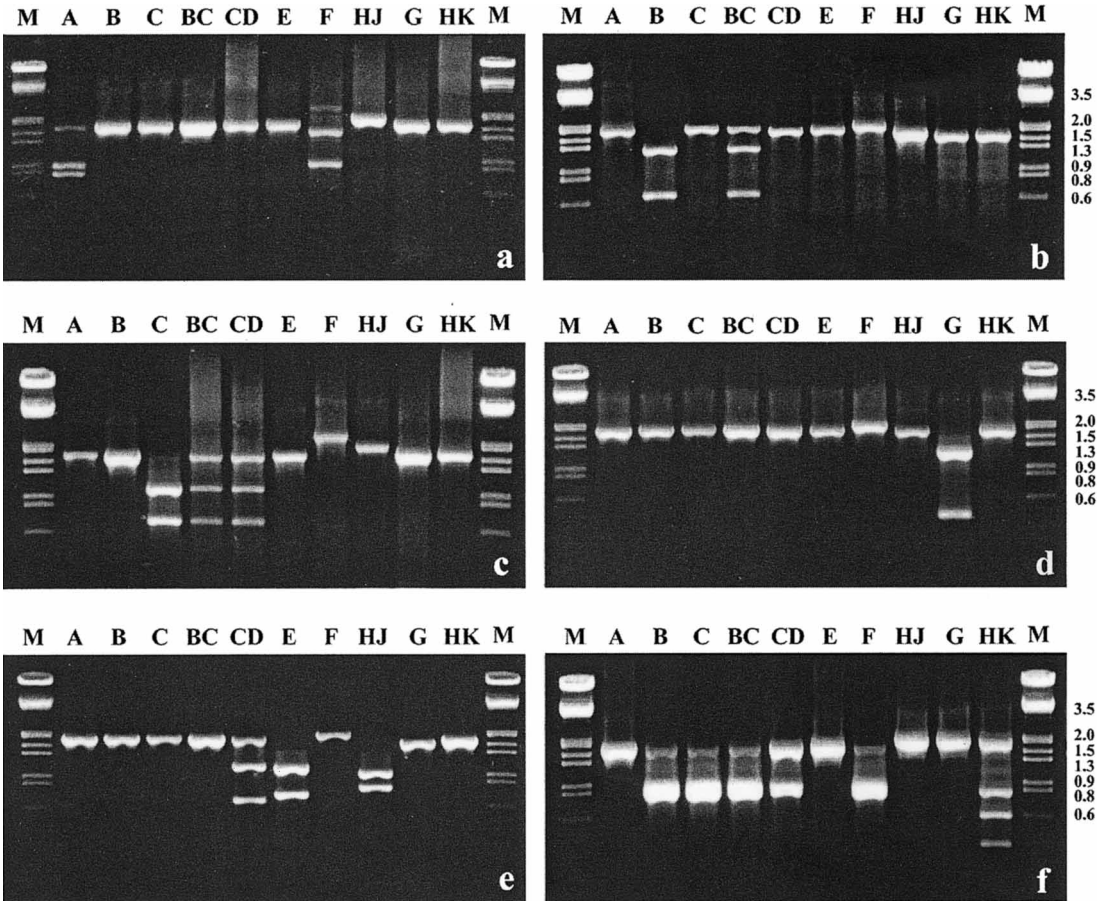


Fig. 2. Restriction profiles of the PCR amplification of the *Adh* gene regions followed by digestion with restriction enzymes. (*a*) *Adh2* + *Sac*II; (*b*) *Adh1* + *Sac*II; (*c*) *Adh2* + *Eco*NI; (*d*) *Adh1* + *Ban*II; (*e*) *Adh1* + *Afl*III; (*f*) *Adh2* + *Bst*EII. Types of genomes are labeled above lanes. The species chosen to represent the genomes are as follows: A, *O. rufipogon*; B, *O. punctata*; C, *O. officinalis*; BC, *O. minuta*; CD, *O. alta*; E, *O. australiensis*; F, *O. brachyantha*; G, *O. granulata*; HJ, *O. ridleyi*; HK, *O. schlechteri*. Lane M is the size marker, and sizes of the fragments (kb) are labeled at the side.



genomes. Recent studies have also pointed out the misidentification of *Oryza* species (Martin et al. 1997; Aggarwal et al. 1999). By using RAPD markers, for example, Martin et al. (1997) showed that five accessions determined as *O. meridionalis* and 4 out of 22 accessions designated as *O. glumaepatula* were misidentified and should be *O. rufipogon*. The results suggest that a certain portion of germplasm in gene banks is misidentified.

Approximately one third of *Oryza* species are tetraploids with different genome constitutions (Khush 1997; Ge et al. 1999). Their taxonomy and identification are always complex and confusing, in particular for those species with the

Table 2. Combinations of two *Adh* genes and restriction enzymes, and their utility in identifying *Oryza* genomes.

No.	Combination	Target genome and its identification	Other genomes
1	*Adh1* + *Tth*111I	A, one band	Two or three bands
2	***Adh2* + *Sac*II**	A and F, two bands (0.80–0.82, 0.89–0.91 for A; 0.83, 1.53 for F)	One band
3	*Adh1* + *Dra*I	A and B, two bands (0.20–0.21, 1.68–1.69 for A; 0.22, 1.70 for B)	One band
4	***Adh1* + *Sac*II**	B, two bands (0.66, 1.25)	One band
5	*Adh1* + *Fau*I	B, two bands (0.66, 1.25)	One band
6	*Adh2* + *Hpa*I	B, two bands (0.42, 1.26)	One band
7	*Adh1* + *Eco*0109I	B and HJ, two bands (0.73, 1.19 for B; 0.25, 1.49 for HJ)	One band
8	*Adh1* + *Ppu*MI	B and HJ, two bands (0.73, 1.19 for B; 0.25, 1.48 for HJ)	One band
9	*Adh2* + *Eco*RI	B, C, and G, two bands (0.19, 1.49 for B; 0.62–0.63, 1.06–1.07 for C; 0.17, 1.52 for G)	One band
10	***Adh2* + *Eco*NI**	C, two bands (0.66, 1.04)	One band
11	***Adh1* + *Afl*III**	D(E) and HJ, two bands (0.62–0.69, 1.11–1.14 for D(E); 0.76, 1.04 for HJ)	One band
12	*Adh1* + *Mme*I	D(E), one band	Two or three bands
13	*Adh1* + *Sap*I	F, two bands (0.63, 1.17)	One band
14	*Adh1* + *Hgi*AI	F, two bands (0.15, 1.65)	One band
15	*Adh2* + *Swa*I	F and HK, two bands (0.37, 1.99 for F; 0.21, 1.72 for HK)	One band
16	***Adh1* + *Ban*II**	G, two bands (0.56, 1.07)	One band
17	*Adh1* + *Avr*II	G, two bands (0.37, 1.26)	One band
18	*Adh2* + *Bst*BI	G, one band	Two bands
19	Adh2 + *Pvu*II	G and HK, one band	Two or three bands
20	Adh1 + *Eco*RI	HJ, two bands (0.23, 1.57)	One band
21	Adh1 + *Bst*EII	HJ, one band	Two or three bands
22	***Adh2* + *Bst*EII**	HK, four bands (0.34, 0.56, 0.79, 1.68)	One or two bands

Note: Restriction sites of those combinations with bold face are shown in Fig. 1, and their restriction patterns are shown in Fig. 2. Numbers in parentheses indicate the reference sizes (kb) of each band.

B, C, BC, and CD genome types (Vaughan 1989). For example, *O. officinalis* (CC) and *O. minuta* (BBCC) are often misidentified with each other, the same being true for the diploid *O. punctata* (BB) and the tetraploid *O. punctata* (BBCC). The recent finding of a tetraploid form of *O. officinalis* causes even more confusion in the identification of these species complexes. Wang et al. (1992) found that one accession with the original name of *O. officinalis* (CC) was actually *O. latifolia* (CCDD). Aggarwal et al. (1999) indicated that one accession identified as *O. malampuzhanensis* (BBCC) contained the CD genome, and two accessions of diploid *O. eichingeri* (CC) were actually tetraploids with the BC genome.

Consequently, misidentification poses a serious problem to the efficient utilization of the wild rice germplasm for breeding and research material. In addition to the traditional cytogenetic methods, recent molecular approaches have been used for genome identification. Tandemly repeated sequences specific to the A genome have already been isolated in rice (Wu and Wu 1987; De Kochko et al. 1991). The same is true of the C (Zhao et al. 1989; Reddy et al. 1993), E, and F genomes (Cordesse et al. 1992). Reddy et al. (1993) found a tandemly repeated DNA sequence that was highly specific for the C genome because it did not hybridize to any of the other genomes. This sequence, however, was absent from some *O. officinalis* (CC) accessions and from the CCDD species, and can not be generally used to identify all C genome species. Similarly, the rDNA units from the C genome have apparently been lost in the CCDD species (Cordesse et al. 1992). Recently, total genomic DNA hybridization and dot hybridization have been used to identify rice genomes (Aggarwal et al. 1997; McIntyre and Winberg 1998). However, both methods involve relatively complex experimental procedures, such as making radioactive probes, and require strict control of experimental conditions, such as washing stringency, to allow the results to be repeatable.

In this study, we identified restriction sites that are specific to each genome by comparing sequences of all of the *Oryza* species for each *Adh* locus. Once the identity of a sample needs to be verified, a simple experimental procedure is designed to meet the need. It includes only three steps: (*i*) amplifying a nuclear locus with the specific PCR primers; (*ii*) digesting PCR products with restriction enzymes that identify a given genome; and (*iii*) running the digested products on an agarose gel and documenting photographically. This technique has a few advantages. First, it produces highly reproducible and predictive restriction profiles because the sequences of the target genes are already known. Second, it is a rapid and inexpensive method involving only PCR equipment and common restriction enzymes. Finally, this technique does not involve complex experimental procedures such as Southern blot hybridization and thus can be adopted in many laboratories with conventional equipment.

It is worthwhile to note that the use of only one restriction enzyme and one amplification is sufficient to identify the genome(s) of a sample in question. However, using more than one combination (Table 2) will further confirm the identification of this sample. We tested additional species of each genome listed in Table 1 to verify the utility of this identification system and found the genome identification accurate for each species.

One problem that remains to be solved is to distinguish between the D and E genomes. As shown in Table 2, these two genomes have similar restriction sites in all of the combinations. Therefore, we treated the D and E genomes as

having an equivalent diagnostic value and put the E in parentheses. Based on our previous phylogenetic study (Ge et al. 1999), species with genomes D and E formed a monophyly on both *Adh1* and *Adh2* trees with high bootstrap support (100%), because the sequences of the D genome species are very similar to that of the E genome species for both *Adh1* and *Adh2*. It is not unexpected, therefore, that no combination of *Adh* genes and restriction enzymes can be found that could distinguish between the D and E genomes efficiently (Table 2).

It is worth mentioning, however, that size variation of restriction fragments occurred in some of the genomes that have more than one species. For example, there existed minor variation in the length of the two bands for three D genome species when treated by the combination *Adh1* + *Afl*III (combination 11 in Table 2). As shown in Fig. 2*e*, two bands exist for the D and E genomes. By surveying all sequences of three CD genome species, we found that the longer band ranged from 1.13 to 1.14 kb, and the shorter band from 0.62 to 0.69 kb. Consequently, it is impossible to tell apart these two genomes efficiently. This is because the sizes of the two restriction fragments of the E genome species fall within the size range of the CD genome species although minor differences in the length of the shorter bands appeared between one of the CD genome species (*O. alta*) and the E genome species (0.63 vs. 0.69 kb) in Fig 2*e*.

There might be questions concerning the general utility and efficiency of the gene-specific PCR–RFLP for genome differentiation and that it could be affected by the sequence polymorphism of *Adh* genes among species with the same genome type, but this concern has been considered in our design. First, in our previous phylogenetic analysis, both *Adh1* and *Adh2* trees showed that all of the sequences from the same genome type formed a monophyletic group with high bootstrap support (100%) (Ge et al. 1999). Second, most of the cutting sites are located in the coding region of the two genes, which are relatively conserved across species (Fig. 1). Finally, we have surveyed the sequences from at least one accession each of all existing species for individual genomes (e.g., eight species for the A genome, three species for the CD genome, etc.) to identify the informative restriction sites. It is less likely that a new sequence with different cutting sites will occur within a genome.

The gene-specific PCR–RFLP technique can be applied to genome determination of unknown materials and detection of true hybrids between species with different genomic constitutions. In addition, it can also be used as a chromosome-specific marker to follow *Adh*-gene located chromosomes in introgression programs. Theoretically, the technique proposed here could extend to species and variety identification as a quick and efficient approach, given that a suitable DNA fragment is chosen and adequate variation exists.

Acknowledgements

We thank the International Rice Research Institute (Manila, Philippines) for providing plant material for this study. This research was supported by National Natural Science Foundation of China (30025005; S.G.) Michigan State University, and the National Science Foundation of the U.S.A., (T.S.) and the Chinese Academy of Sciences (KZ951-B1-102).

References

Aggarwal, R.K., Brar, D.S., and Khush, G.S. 1997. Two new genomes in the *Oryza* complex identified on the basis of molecular divergence analysis using total genomic DNA hybridization. Mol. Gen. Genet. **254**: 1–12.

Aggarwal, R.K., Brar, D.S., Nandi, S., Huang, N., and Khush, G.S. 1999. Phylogenetic relationships among *Oryza* species revealed by AFLP markers. Theor. Appl. Genet. **98**: 1320–1328.

Arnold, M.L., Buckner, C.M., and Robinson, J.J. 1991. Pollen-mediated introgression and hybrid speciation in Louisiana irises. Proc. Natl. Acad. Sci. U.S.A. **88**: 1398–1402.

Badenes, M.L., and Parfitt, D.E. 1995. Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. Theor. Appl. Genet. **90**: 1035–1041.

Cordesse, F., Grellet, F., Reddy, A.S., and Delseny, M. 1992. Genome specificity of rDNA spacer fragments from *Oryza sativa* L. Theor. Appl. Genet. **83**: 864–870.

De Kochko, A., Kiefer, M.C., Cordesse, F., Reddy, A.S., and Delseny, M. 1991. Distribution and organisation of a tandemly repeated 352 bp sequence in the *Oryza* family. Theor. Appl. Genet. **82**: 57–64.

Doyle, J.J., and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**: 11–15.

Ge, S., Sang, T., Lu, B.R., and Hong, D.Y. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc. Natl. Acad. Sci. U.S.A. **96**: 14400–14405.

Isshiki, S., Uchiyama, T., Tashiro, Y., and Miyazaki, S. 1998. RFLP analysis of a PCR amplified region of chloroplast DNA in eggplant and related *Solanum* species. Euphytica, **102**: 295–299.

Jena, K.K., and Kochert, G. 1991. Restriction fragment length polymorphism analysis of CCDD genome species of the genus *Oryza* L. Plant Mol. Biol. **16**: 831–839.

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

Liston, A. 1992. Variation in the chloroplast genes *proc1* and *proc2* of the genus *Astragalus* (Fabaceae): evidence from restriction site mapping of a PCR-amplified fragment. Am. J. Bot. **79**: 953–961.

Lu, B.R., Naredo, M.B.E., Juliano, A.B., and Jackson, M.T. 2000. Preliminary studies on taxonomy and biosystematics of the AA genome *Oryza* species (Poaceae). *In* Grasses: systematics and evolution. *Edited by* S.W.L. Jacobs and J. Everett. CSIRO, Melbourne. pp. 51–58.

Marck, C. 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16**: 1829–1836.

Martin, C., Juliano, A., Newbury, H.J., Lu, B.R., and Jackson, M.T. 1997. The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. Genet. Res. Crop Evol. **44**: 175–183.

McIntyre, C.L., and Winberg, B.C. 1998. A rapid means of identifying wild rice species DNA using dot blots and genome-specific rDNA probes. Genome, **41**: 391–395.

Nayar, N.M. 1973. Origin and cytogenetics of rice. Adv. Genet. **17**: 153–292.

Reddy, A.S., Kiefer-Meyer, M.C., and Delseny, M. 1993. Characterization of new variants of a satellite DNA from *Oryza officinalis*, specific for the CC genome of wild rice. Genome, **36**: 750–761.

Sarkar, R., and Raina, S.N. 1992. Assessment of genome relationships in the genus *Oryza* L. based on seed-protein profile analysis. Theor. Appl. Genet. **85**: 127–132.

Tanksley, S.D., and McCouch, S.R. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Science (Washington, D.C.), **277**: 1063–1066.

Transue, D.K., Fairbanks, D.J., Robison, L.R., and Andersen, W.R. 1994. Species identification by RAPD analysis of grain amaranth genetic resources. Crop Sci. **34**: 1385–1389.

Vaughan, D.A. 1989. The genus *Oryza* L.: current status of taxonomy. International Rice Research Institute, Manila, Philippines.

Vaughan, D.A. 1994. The wild relatives of rice: a genetic resources handbook. International Rice Research Institute, Manila, Philippines.

Virk, P.S., Ford-Lloyd, B.V., Jackson, M.T., and Newbury, H.J. 1995. Use of RAPD for the study of diversity within plant germplasm collections. Heredity, **74**: 170–179.

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 nuclear RFLPs. Theor. Appl. Genet. **83**: 565–581.

Wu, T.Y., and Wu, R. 1987. A new rice repetitive DNA shows sequence homology to both 5S RNA and tRNA. Nucleic Acids Res. **15**: 5913–5923.

Zhao, X., Xie, Y., and Wu, R. 1989. Genome specific repetitive sequences in the genus *Oryza*. Theor. Appl. Genet. **78**: 201–209.