

• Rapid Communication •

Identification of *Oryza* Species with the CD Genome Based on RFLP Analysis of Nuclear Ribosomal ITS Sequences

BAO Ying^{1,2}, GE Song^{1*}

(1. Laboratory of Systematic and Evolutionary Botany, Institute of Botany, The Chinese Academy of Sciences, Beijing 100093, China;

2. Department of Biology, Qufu Normal University, Qufu 213765, China)

Abstract: The genus *Oryza* consists of approximately 24 species with 10 recognized genome types (A, B, C, BC, CD, E, F, G, HJ, HK). The species with the CD genome are endemic to the Central and South America and comprise three species, i.e. *O. latifolia* Desv., *O. alta* Swallen and *O. grandiglumis* (Doell) Prod. Although the three species can be easily distinguished from the species with the other genomes in *Oryza* by their morphological characteristics, it is very difficult to distinguish them from one another. Recent studies suggested that *O. alta* and *O. grandiglumis* should be the same species (*O. grandiglumis*), whereas *O. latifolia* retains its taxonomic position unchangeably. In this paper, a total of 77 clones of the nuclear ribosomal internal transcribed spacer (ITS) from 11 samples representing different geographical races were sequenced. Using DNA Strider 1.2 software, the restriction enzyme digestion sites of these clone sequences were analyzed. Based on restriction fragment length polymorphism (RFLP) of ITS sequences, a method to identify the species with CD genome in *Oryza* was proposed. The method is rapid and convenient and all experimental procedure includes only three steps: (1) to amplify the ITS fragment with the routine primers; (2) to digest PCR products with restriction enzymes *Fok* and *Dra* or both; (3) to run the digested product on 1% agarose gel and identify the sample based on the restriction profiles.

Key words: *Oryza*; identification; CD genome; RFLP; ITS

The genus *Oryza* consists of approximately 24 species with 10 recognized genome types (A, B, C, BC, CD, E, F, G, HJ, HK) (Ge *et al.*, 1999). The species with the CD genome are endemic to the Central and South America and comprise three species, i.e. *O. latifolia*, *O. alta* and *O. grandiglumis* (Vaughan, 1989). These species are perennial with erect and higher plants as well as big and open panicles (Nayar, 1973). Of them, *O. latifolia* is widely distributed in Central and South America and on the Caribbean islands. In contrast, *O. alta* and *O. grandiglumis* occur only in South America (Vaughan, 1989).

In the genus *Oryza*, the species with the CD genome are easily distinguished from the species with other genome types by their larger plants, broader leaves (more than 2 cm) and unique geographical distribution (Vaughan, 1989). However, it is difficult to distinguish these three species from one another. Up to date, different opinions have been raised regarding whether the three species are treated as separate or a single species (Nayar, 1973; Jena and Kochert, 1991; Buso *et al.*, 2001). Many molecular studies using the nuclear and chloroplast gene sequences (Ge

et al., 1999), AFLP markers (Aggarwal *et al.*, 1999) and RFLP of repetitive sequences (Federici *et al.*, 2002) all indicated that *O. alta* and *O. grandiglumis* were closer to each other than either to *O. latifolia*. Recently, we studied the CD genome species based on the sequences from two nuclear genes (*Adh1* and *Adh2*), two chloroplast genes (*matK* and *trnL-trnF*) and the nuclear ribosomal internal transcribed spacer (ITS), and found that the accessions sampled from *O. alta* and *O. grandiglumis* were always clustered together and separated clearly from *O. latifolia* on multiple-gene trees (unpublished). Therefore, we suggest that *O. alta* and *O. grandiglumis* should be treated as same species (*O. grandiglumis*), while *O. latifolia* remains a separate species.

A correct classification is the premise for effective utilization of plant resources and germplasm, especially for wild rice which contains an enormous gene pool for the genetic improvement of rice cultivars (Vaughan, 1989; Ge *et al.*, 1999). Here we present a rapid and convenient method to distinguish the species with the CD genome based on the RFLP profile of the nuclear ribosomal internal transcribed spacer (ITS). This method can facilitate identification of *Oryza*

Received: 2002-11-07 Accepted: 2003-03-17

Supported by the National Natural Science Foundation of China (30025005) and The Chinese Academy of Sciences (KSCXZSW-101A).

*Author for correspondence. E-mail: <gesong@ns.ibcas.ac.cn>.

species with the CD genome, and can be a potential way for utilization in the wild rice germplasm collection.

1 Materials and Methods

A total of 11 accessions were sampled in the study, including three accessions from each of *Oryza alta* Swallen and *O. grandiglumis* (Doell) Prod., five from *O. latifolia* Desv. The species names, accession numbers and origins are listed in Table 1. Plants were grown from seeds kindly provided by the International Rice Genebank at the International Rice Research Institute (IRRI) in the Philippines and developed in the greenhouse.

Total DNA was isolated from fresh leaves of plants individually by using CTAB method (Doyle and Doyle, 1987). Nuclear DNA ITS region was amplified with primer ITS1 (5'-AGAAGTCGTAACAAGGTTTCCGTAGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') (White *et al*, 1990). Each amplification reaction (25 μ L final volume) contains 10 - 20 ng template DNA, 10 mmol/L Tris-HCl (pH 8.3), 2.0 mmol/L MgCl₂, 0.2 mmol/L of each dNTP, 5 pm each primer, and 0.75 units of *Taq* polymerase. In addition, 8% dimethylsulfoxide (DMSO) was added to the reaction mixture to facilitate denaturation during amplification (Baldwin *et al*, 1995). The PCR procedure includes 30 cycles of 1 min denaturation at 94 °C, 1 min annealing at 52 °C, and 1.5 min extension at 72 °C and a final 10-min extension at 72 °C. Amplifications were performed in PTC-200 (PE) thermocycler.

We have cloned and sequenced the internal transcribed spacer (ITS) for the CD genome species, including 14 clones for *O. alta*, seven clones for *O. grandiglumis* and 56 clones for *O. latifolia*. The ITS sequences of these species were deposited in the GenBank database under the accession numbers listed in Table 1. In order to obtain the diagnostic restriction profiles, we surveyed the restriction sites of the ITS sequences using DNA Strider 1.2 (Marck, 1988). The strategy for determining species-specific sites followed that of Ge *et al* (2001).

For the purpose of confirmation of the diagnostic sites, 2 to 5 μ L of PCR products of 11 accessions representing different geographical races were digested in each 10 μ L reaction containing 1 U of restriction enzyme. Digested PCR products were electrophoresed on 1% TBE agarose gels. The gels were stained with ethidium bromide, visualized, and photographed under UV light.

2 Results and Discussion

Using DNA Strider, we surveyed 170 restriction enzymes for the restriction sites of the ITS sequences of the CD genome species. Of them, the enzymes that had restriction sites conserved across accessions of the same species but variable among different species were chosen for the purpose of identifying species. So, two restriction enzymes of *Fok* I and *Dra* I were chosen by which we could distinguish *O. latifolia* from *O. grandiglumis* (*O. alta*) unambiguously. However, no restriction site was found to distinguish *O. grandiglumis* from *O. alta*, demonstrating close relationship between them as indicated by many previous studies (Aggarwal *et al*, 1999; Ge *et al*, 1999; Federici *et al*, 2002).

Restriction enzyme *Fok* I has no cutting site in the ITS region of *O. latifolia* but one cutting site in the middle of the ITS regions of *O. alta* and *O. grandiglumis*, resulting in a 700-bp band for *O. latifolia* and two nearly equal bands (~350 bp) for *O. alta* and *O. grandiglumis*. For restriction enzyme *Dra* I, on the contrary, one cutting site in *O. latifolia* but no cutting site in *O. alta* and *O. grandiglumis* were observed, resulting two bands (200 bp and 500 bp, respectively) for *O. latifolia* and one band (700 bp) for *O. alta* and *O. grandiglumis*. The restriction profiles of the two enzymes are shown in Fig. 1. Consequently, we can distinguish *O. latifolia* from *O. alta* or *O. grandiglumis* clearly based on the restriction digestions of PCR-amplified ITS sequences by *Dra* I and *Fok* I.

In summary, the method proposed in the present study

Table 1 List of samples

Species	Accession	Origin	Genbank number
<i>Oryza alta</i>	100161	Brazil	AF513601-AF513606
<i>O. alta</i>	100967	Surinam	
<i>O. alta</i>	105143	Guyana	AF514909-AF514917
<i>O. grandiglumis</i>	105664	Brazil	AF520780, AY151368
<i>O. grandiglumis</i>	105669	Brazil	AY151363-AY151367
<i>O. grandiglumis</i>	101405	Brazil	
<i>O. latifolia</i>	100167	Costa Rica	AF513553-AF513564
<i>O. latifolia</i>	100171	Guatemala	AF513565-AF513576
<i>O. latifolia</i>	100914	Mexico	AF513577-AF513588
<i>O. latifolia</i>	105141	Costa Rica	AY151355-AY151362
<i>O. latifolia</i>	105145	Colombia	AF513589-AF513600

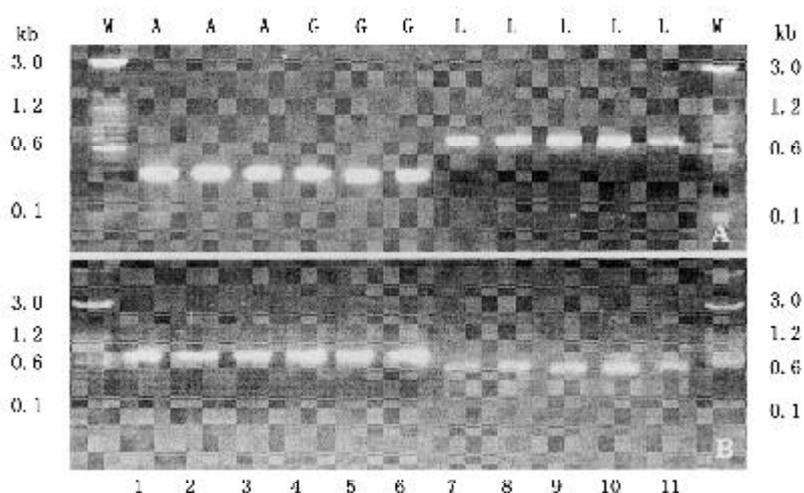


Fig.1. Restriction profiles of PCR-amplified ITS regions followed by digestion with restriction enzymes. **A.** Digestion profiles by restriction enzyme *Fok* I. **B.** Digestion profiles by restriction enzyme *Dra* I. Lanes 1-3, *Oryza alta*; Lanes 4-6, *O. grandiglumis*; Lanes 7-11, *O. latifolia*.

is rapid and convenient. The experimental procedure includes only three steps: (1) to amplify the ITS fragment with the routine ITS primers; (2) to digest PCR products with restriction enzymes *Fok* I and *Dra* I or both; and (3) to run the digested product on 1% agarose gel and identify the sample based on the restriction profiles. The method can exclude the influence by artificial factors and improve validity of identification greatly. More importantly, this method offers an attempt to distinguish these species at any stage of their life cycle.

References:

- Aggarwal R K, Brar D S, Nandi S, Huang N, Khush G S. 1999. Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theor Appl Genet*, **98**:1320-1328.
- Baldwin B G, Sanderson M J, Wojciechowski M F, Campbell C S, Donghue M J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann Missori Bot Gard*, **82**:247-277.
- Buso G S C, Rangel P H, Ferreira M E. 2001. Analysis of random and specific sequences of nuclear and cytoplasmic DNA in diploid and tetraploid American wild rice species (*Oryza* spp.). *Genome*, **44**:476-494.
- Doyle J J, Doyle J L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*, **19**:11-15.
- Federici M T, Shcherban A B, Capdevielle F, Francis M, Vaughan D. 2002. Analysis of genetic diversity in the *Oryza officinalis* complex. *Biotechnology*, **5**:173-181.
- Ge S, Sang T, Lu B R, Hong D Y. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proc Natl Acad Sci USA*, **96**:14400-14405.
- Ge S, Sang T, Lu B R, Hong D Y. 2001. Rapid and reliable identification of rice genomes by RFLP analysis of PCR-amplification *Adh* genes. *Genome*, **44**:1136-1142.
- Jena K K, Kochert G. 1991. Restriction fragment length polymorphism analysis of CCDD genome species of the genus *Oryza* L. *Plant Mol Biol*, **16**:831-839.
- 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.
- Nayar N M. 1973. Origin and cytogenetics of rice. Caspari E W. *Advances in Genetics*. New York: Academic Press. 217-220.
- White T J, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Innis M, Gelfand D, Sninsky J, White T. *PCR Protocols: a Guide to Methods and Applications*. San Diego: Academic Press. 315-322.
- Vaughan D A. 1989. The genus *Oryza* L. Current status of taxonomy. Manila: *IRRI Res Paper Series*, **138**:9.

(Managing editor: HE Ping)

根据核 DNA 的 ITS 序列的 RFLP 分析鉴定稻属 CD 染色体组物种

包颖^{1,2} 葛颂^{1*}

(1 .中国科学院植物研究所系统与进化植物学重点实验室,北京 100093 ;

2 .山东曲阜师范大学生物系,曲阜 213765)

摘要 : 稻属物种的染色体组类型有10种,其中具CD染色体组的物种有3种(*Oryza alta* Swallen, *O. grandiglumis* (Doell) Prod.和*O. latifolia* Desv.),仅分布在美洲中部和南部。这3个种在形态上和稻属其他染色体组的种比较容易区别,但它们彼此之间鉴别比较困难。近年来的研究表明,*O. alta*和*O. grandiglumis*应归并为1个种(*O. grandiglumis*),而*O. latifolia*应保持不变。本文基于代表不同分布区11个样品的核糖体转录间隔区(ITS)的77个克隆序列数据,利用DNA Strider1.2软件进行了限制性酶切位点分析,提出了一个鉴别CD染色体组物种的方法。方法的具体步骤是:(1)用通用引物扩增ITS片段;(2)用特异性的限制性内切酶Fok I 和/或Dra 消化PCR扩增产物;(3)用1%的琼脂糖电泳并根据消化产物的多态性特征来鉴别不同物种。基于本文提出的核糖体ITS限制性片段多态性(RLFP)分析,可以快速和可靠地将稻属CD染色体组物种区别开来。

关键词 : 稻属; *Oryza alta*; *O. grandiglumis*; *O. latifolia*; 核糖体ITS; RFLP

中图分类号: Q943 文献标识码: A 文章编号: 0577-7496(2003)07-0762-04

收稿日期:2002-11-07 接收日期:2003-03-17

基金项目:国家自然科学基金(30025005);中国科学院知识创新工程重要方向项目(KSCXZSW-101A)。

*通讯作者。E-mail: <gesong@ns.ibcas.ac.cn>.

(责任编辑: 贺 萍)