J.Sci.Soc.Thailand, 22 (1996) 89-94

PARTIAL cDNA SEQUENCE OF CHLOROPLAST 23S RIBOSOMAL RNA FROM *INDICA* RICE.

PONGSOPEE ATTASART [1,2] AND BURACHAI SONTHAYANON [1,2]

[1] Department of Biochemistry, Faculty of Science, Mahidol University, Rama VI Rd., Bangkok 10400.

[2] Institute of Science and Technology for Research and Development, Mahidol University, Salaya, Phutthamonthon District, Nakorn Pathom 73140, Thailand.

(Received February 5, 1996)

ABSTRACT

A partial cDNA sequence of rrn23 *of* indica *rice (*Oryza sativa *L. subspecies* indica *variety Khao Dawk Mali 105) was obtained. The 1130 bp sequence, derived and confirmed from three independently isolated cDNA clones, covered nucleotide positions 1760 to 2887 in domains IV and V of the 23S ribosomal RNA sequence of rice. The newly characterized indica sequence revealed a 99.4 % sequence identity to the corresponding region of a* japonica *rice chloroplast DNA and a 99.0 % sequence identity to that of maize. The differences found were present in both domains which are normally regarded as having no apparent secondary structure. The newly characterized sequence data would be useful in varietal cataloging of cultivated rice.*

INTRODUCTION

Cultivated rice (*Oryza sativa* L.) is one of the world's most important crops, with over 50 % of the population consuming rice as their staple diet[1]. Of its two major subspecies, *japonica* and *indica*, the *indica* is widely grown in tropical areas while *japonica* is grown mainly in temperate areas[2]. For *indica* rice, thousands of varieties are planted in various localities worldwide. Most of the *indica* varieties are morphologically similar but the shapes of plant and grain are generally distinguishable from *japonica*[2]. However, some varieties are not easily categorized as either subspecies based on morphologies alone. Additional genetic markers, such as isozyme and RFLP markers, are needed to put a number of varieties into the correct taxa[3,4] .

Ribosomal RNA genes have been well characterized in hundreds of plants, animals, and prokaryotic organisms[5]. Their sequences are useful in investigation of protein synthesis mechanism[6-13] as well as in taxonomy and phylogenetics of living organisms[14-21]. For cultivated rice, the sequence of *rrn16, rrn18, rrn23* , and *rrn26* of the subspecies *japonica* have been reported[22,23]. However, the *rrn23* sequence of the predominant subspecies *indica* has not been determined and it is generally assumed that those of the two subspecies are identical. In this paper, we report a partial sequence of a chloroplast 23S rRNA based on cloned cDNA sequences which revealed a number of sequence differences from that of the

GenBank accession number U49840

japonica rice in domains IV and V. Our finding on these differences opens the possibility of using the *rrn23* sequence as another marker for the cataloging of the two rice subspecies.

MATERIALS AND METHODS

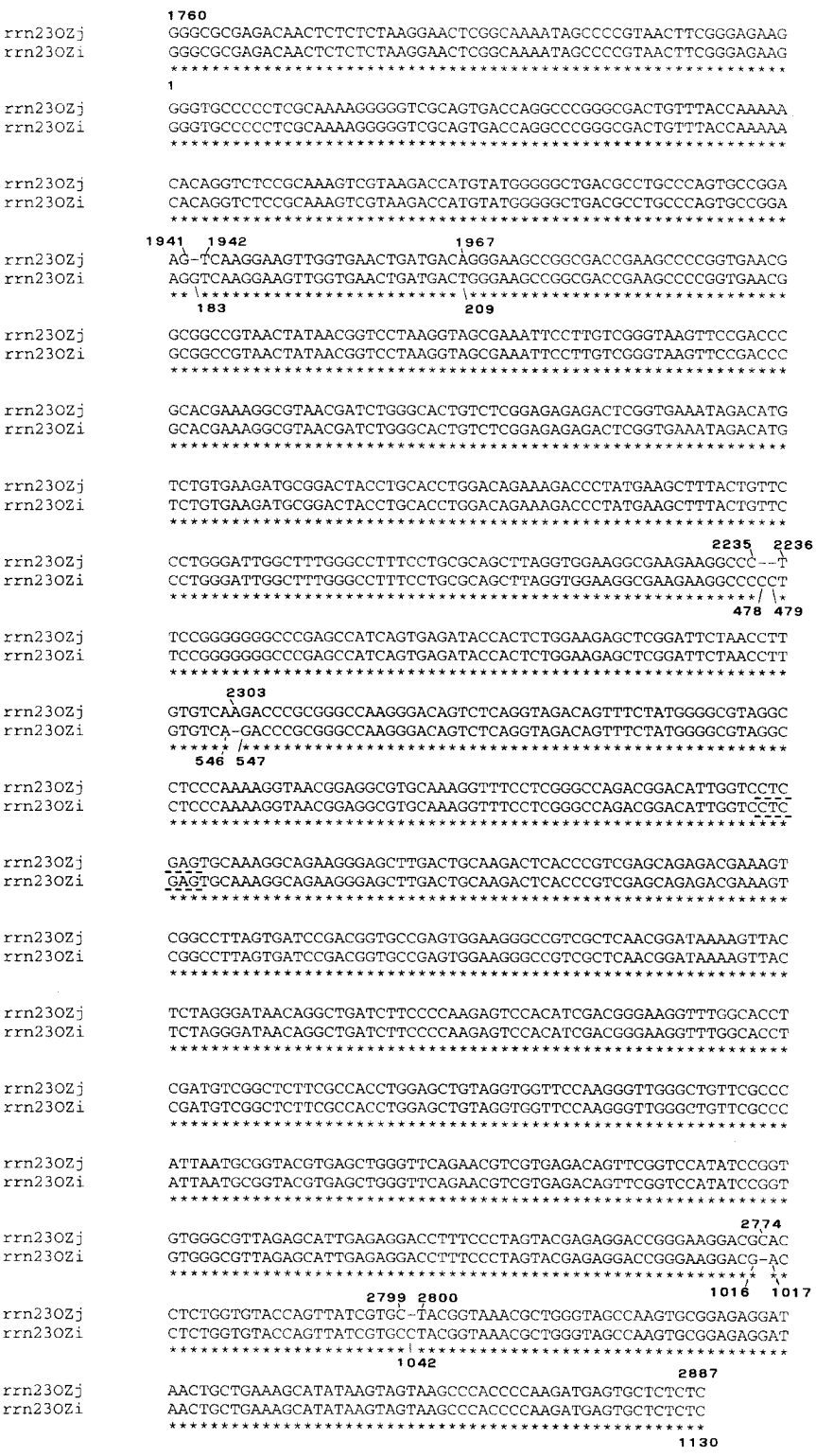
A number of cDNA clones were isolated from cDNA library of an *indica* rice (wounded and UV-induced), Khao Dawk Mali 105 variety, constructed in Lambda ZAP II vector (Promega) and *E. coli* XL1-Blue host (kindly provided by Dr. Skorn Mongkolsuk of Department of Biotechnology, Faculty of Science, Mahidol University). They were screened by *in situ* hybridization using ^{32}P-labeled 3.7 kb Xho I fragment containing a putative large rRNA gene of *indica* rice mitochondria earlier characterized in our laboratory[24]. Standard Molecular biology techniques were used in this study[25-28] and reagents were either of analytical or Molecular biology grade. Three cDNA clones of different sizes were isolated, purified, and sequenced manually in both strands by Sanger's dideoxy procedure using [α-^{35}S] dATP labelling [29,30]. The sequence of one clone was also reconfirmed by automated DNA sequencing method (Perkin Elmer ABI PRISM^{tm} 377) using dye terminator chemistry. Sequences of all 3 clones were combined and verified to provide a composite sequence. DNA sequence comparison was made by using Clustal V program[31,32] with DNA sequences available from GenBank via electronic mails[33].

RESULTS AND DISCUSSION

The cDNA sequence assembled was derived from 3 cDNA clones from a cDNA library of the well known aromatic *indica* variety of Thailand, Khao Dawk Mali 105 (also known as Jasmine rice). The cDNA clones were screened from a cDNA library prepared from leaf mRNA of stressed and damaged -induced rice plants. After several attempts of library screening, only positive clones with short inserts (< 1 kb) could be isolated. After the cDNA sequence was obtained, comparison was made with known sequences in the database and the identity was assigned as *rrn23* based on high percent sequence identity with those of *E. coli* [34] (around 71%) and *Oryza sativa* chloroplast (99.4%). In fig.1, the composite sequence (called rrn23ozi) derived from our cDNA clones is shown aligned with that of the *japonica* rice chloroplast DNA sequence (labelled as rrn23ozj). From the alignment, it could be seen that the region analyzed corresponded to the 3' part of the *rrn23* sequence.

Fig 1. Partial nucleotide (cDNA) sequence of 23S RNA from chloroplast of *Oryza sativa* L., subspecies *indica*, variety Khao Dawk Mali 105 aligned with that of its counterpart from *japonica* subspecies.

The rrn23ozi DNA is shown in the lower line of the alignment. The sequence from nucleotide position 1 to an internal Xho I site was from one cDNA clone and the rest was from two other clones. The position 1 of this sequence corresponds to position 1760 of a 23S RNA gene from *japonica* rice chloroplast. Asterisks (*) show identical nucleotides between the two sequences. The differences are indicated by the lack of the asterisk. Gaps in the alignment due to the lack of some nucleotides in a sequence are shown by dashes (-).


1760
rrn23OZj GGGCGCGAGACAACTCTCTCTAAGGAACTCGGCAAAATAGCCCCGTAACTTCGGGAGAAG
rrn23OZi GGGCGCGAGACAACTCTCTCTAAGGAACTCGGCAAAATAGCCCCGTAACTTCGGGAGAAG
1
rrn23OZj GGGTGCCCCCTCGCAAAAGGGGGTCGCAGTGACCAGGCCCGGGCGACTGTTTACCAAAAA
rrn23OZi GGGTGCCCCCTCGCAAAAGGGGGTCGCAGTGACCAGGCCCGGGCGACTGTTTACCAAAAA
rrn23OZj CACAGGTCTCCGCAAAGTCGTAAGACCATGTATGGGGGCTGACGCCTGCCCAGTGCCGGA
rrn23OZi CACAGGTCTCCGCAAAGTCGTAAGACCATGTATGGGGGCTGACGCCTGCCCAGTGCCGGA
1941 1942 1967
rrn23OZj AG-TCAAGGAAGTTGGTGAACTGATGACAGGGAAGCCGGCGACCGAAGCCCCGGTGAACG
rrn23OZi AGGTCAAGGAAGTTGGTGAACTGATGACTGGGAAGCCGGCGACCGAAGCCCCGGTGAACG
183 209
rrn23OZj GCGGCCGTAACTATAACGGTCCTAAGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCC
rrn23OZi GCGGCCGTAACTATAACGGTCCTAAGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCC
rrn23OZj GCACGAAAGGCGTAACGATCTGGGCACTGTCTCGGAGAGAGACTCGGTGAAATAGACATG
rrn23OZi GCACGAAAGGCGTAACGATCTGGGCACTGTCTCGGAGAGAGACTCGGTGAAATAGACATG
rrn23OZj TCTGTGAAGATGCGGACTACCTGCACCTGGACAGAAAGACCCTATGAAGCTTTACTGTTC
rrn23OZi TCTGTGAAGATGCGGACTACCTGCACCTGGACAGAAAGACCCTATGAAGCTTTACTGTTC
2235 2236
rrn23OZj CCTGGGATTGGCTTTGGGCCTTTCCTGCGCAGCTTAGGTGGAAGGCGAAGAAGGCCC--T
rrn23OZi CCTGGGATTGGCTTTGGGCCTTTCCTGCGCAGCTTAGGTGGAAGGCGAAGAAGGCCCCCT
478 479
rrn23OZj TCCGGGGGGGCCCGAGCCATCAGTGAGATACCACTCTGGAAGAGCTCGGATTCTAACCTT
rrn23OZi TCCGGGGGGGCCCGAGCCATCAGTGAGATACCACTCTGGAAGAGCTCGGATTCTAACCTT
2303
rrn23OZj GTGTCAAGACCCGCGGGCCAAGGGACAGTCTCAGGTAGACAGTTTCTATGGGGCGTAGGC
rrn23OZi GTGTCA-GACCCGCGGGCCAAGGGACAGTCTCAGGTAGACAGTTTCTATGGGGCGTAGGC
546 547
rrn23OZj CTCCCAAAAGGTAACGGAGGCGTGCAAAGGTTTCCTCGGGCCAGACGGACATTGGTCCTC
rrn23OZi CTCCCAAAAGGTAACGGAGGCGTGCAAAGGTTTCCTCGGGCCAGACGGACATTGGTCCTC
rrn23OZj GAGTGCAAAGGCAGAAGGGAGCTTGACTGCAAGACTCACCCGTCGAGCAGAGACGAAAGT
rrn23OZi GAGTGCAAAGGCAGAAGGGAGCTTGACTGCAAGACTCACCCGTCGAGCAGAGACGAAAGT
rrn23OZj CGGCCTTAGTGATCCGACGGTGCCGAGTGGAAGGGCCGTCGCTCAACGGATAAAAGTTAC
rrn23OZi CGGCCTTAGTGATCCGACGGTGCCGAGTGGAAGGGCCGTCGCTCAACGGATAAAAGTTAC
rrn23OZj TCTAGGGATAACAGGCTGATCTTCCCCAAGAGTCCACATCGACGGGAAGGTTTGGCACCT
rrn23OZi TCTAGGGATAACAGGCTGATCTTCCCCAAGAGTCCACATCGACGGGAAGGTTTGGCACCT
rrn23OZj CGATGTCGGCTCTTCGCCACCTGGAGCTGTAGGTGGTTCCAAGGGTTGGGCTGTTCGCCC
rrn23OZi CGATGTCGGCTCTTCGCCACCTGGAGCTGTAGGTGGTTCCAAGGGTTGGGCTGTTCGCCC
rrn23OZj ATTAATGCGGTACGTGAGCTGGGTTCAGAACGTCGTGAGACAGTTCGGTCCATATCCGGT
rrn23OZi ATTAATGCGGTACGTGAGCTGGGTTCAGAACGTCGTGAGACAGTTCGGTCCATATCCGGT
2774
rrn23OZj GTGGGCGTTAGAGCATTGAGAGGACCTTTCCCTAGTACGAGAGGACCGGGAAGGACGCAC
rrn23OZi GTGGGCGTTAGAGCATTGAGAGGACCTTTCCCTAGTACGAGAGGACCGGGAAGGACG-AC
1016 1017
2799 2800
rrn23OZj CTCTGGTGTACCAGTTATCGTGC-TACGGTAAACGCTGGGTAGCCAAGTGCGGAGAGGAT
rrn23OZi CTCTGGTGTACCAGTTATCGTGCCTACGGTAAACGCTGGGTAGCCAAGTGCGGAGAGGAT
1042
2887
rrn23OZj AACTGCTGAAAGCATATAAGTAGTAAGCCCACCCCAAGATGAGTGCTCTCTC
rrn23OZi AACTGCTGAAAGCATATAAGTAGTAAGCCCACCCCAAGATGAGTGCTCTCTC
1130

The 1130 bp sequence characterized corresponded to nucleotide position 1760 to position 2887 of the *japonica rrn23* sequence[23]. When comparison was made between the sequence with the corresponding region of *rrn23* of the *japonica* rice by Clustal V program, around 99.4 % sequence identity was found. Differences found included a base change (nucleotide 209, which corresponded to the *japonica* sequence position 1967) and a number of deletions or insertions of bases. These included the presence of extra nucleotides for the *indica* sequence at positions 183, 478, 479 and 1042, which corresponded to *japonica* sequence between positions 1941-1942, 2235-2236 and 2799-2800, as well as missing nucleotides of the *indica* sequence positions between 546-547, and 1016-1017, which corresponded to the *japonica* sequence at positions 2303 and 2774, respectively. All of the discrepancies were located in non-hydrogen bonded domains IV and V of the 23S rRNA thus did not affect the predicted secondary structure of the rRNA.

It is quite surprising to see that a gene which is normally regarded as highly conserved across the wide range of organisms in various Kingdoms showed such a degree of sequence differences at the subspecies level, especially in the domain V of 23S rRNA which has been known to have peptidyltransferase activity of the ribosome [6-13]. Perhaps, one could speculate that the mutated nucleotides might not play a role in such catalytic activity. When the DNA sequence was compared to the corresponding region of the 23S RNA sequence of maize[35] which belongs to the same Poaceae (Gramineae) family, a 99.0 % sequence identity was found. Thus the 99.4 % sequence identity found between the two *rrn23* sequences of the two rice subspecies is thus unexpectedly low. However, the fact that rice is a man-made crop as well as being self-fertilizing could explain the relatively high degree of divergence between descendents of this highly conserved gene.

The sequence divergence found between the two subspecies would be useful for phylogeneticists. Since cultivated rice is a man-made crop, the differences would be a useful parameter for following changes created by domestication and human-selection during the past few millennia of rice breeding. Furthermore, the divergent nucleotide sequences are also potentially useful as another criterion in cataloging the thousands of major rice varieties into the two subspecies for plant breeders' germplasm. Rice breeders frequently have problems in telling the two subspecies apart based on the two subspecies' general morphology. Although RFLP and RAPD analysis have been recently introduced to help in such varietal cataloging [36-38], the availability of the two rice chloroplast 23S rRNA sequences opens the possibility of using PCR sequencing to directly and rapidly analyze the chloroplastic rRNA gene.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Skorn Mongkolsuk for providing the cDNA library for our work. Research in the second author (BS) 's laboratory was supported by the Rockefeller Foundation (RF) and Thailand's National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency (NSTDA), Ministry of Science, Technology and Environment, to whom we are indebted. BS was a recipient of RF's Biotechnology Career Fellowship and PA was a recipient of NSTDA's graduate fellowship. General and financial support by Faculty of Science, Mahidol University to BS are also appreciated. The authors would like to thank Dr. Prapon Wilairat for kindly reading the manuscript and providing several useful comments and suggestions.

REFERENCES

1. Kochert G. (1992) Rice as a model system. In *Grass Evolution and Domestication* (Chapman,G.P., ed.), Cambridge University Press, pp. 290-315.
2. Hoshikawa, K. (1989) The Growing Rice Plant : An anatomical monograph. First English edition. Nosan Gyosan Bunka Kyokai, Tokyo.
3. Glaszmann, J.C. (1985) A varietal classification of Asian cultivated rice (*Oryza sativa* L.) based on isozyme polymorphism. In Rice Genetics. International Rice Research Institute, Manila, Philippines, pp.1-85.
4. Ishii,T., and K. Tsunewaki. (1991) Chloroplast genome differentiation in Asian cultivated rice. Genome **34**, 818-826
5. Gutell, R.R., Gray, M.W., Schnare, M.N. (1993) A Compilation of Large Subunit (23S Like and 23S Like) Ribosomal RNA Structures - 1993. Nucleic Acid Res. **21**, 3055-3074
6. Noller, H.F. (1991) Ribosomal RNA and translation. Ann. Rev. Bioch. 60, 191-227
7. Gutell, R.R. Larsen, N. Woese, C.R. (1994) Lessons from an Evolving Ribosomal RNA - 16S and 23S Ribosomal RNA Structures from a Comparative Perspective. Microb. Rev. **58**, 10-26
8. Aagaard, C. and Douthwaite, S. (1994) Requirement for a Conserved, Tertiary Interaction in the Core of 23S Ribosomal RNA. Proc. Natl. Acad. Sci. USA **91**, 2989-2993
9. Kooi, E.A., Rutgers, C.A., Mulder, A., Vantriet, J., Venema, J., Raue, H.A. (1993) The Phylogenetically Conserved Doublet Tertiary Interaction in Domain III of the Large Subunit Ribosomal RNA Is Crucial for Ribosomal Protein Binding. Proc. Natl. Acad. Sci. USA **90**, 213-216
10. Lieberman, K.R., and A.E. Dahlberg. (1995) Ribosome-catalyzed peptide-bond formation. Prog. in Nucleic Acid *Res. and Mol. Biol.* **50**, 1-23
11. Moazed, D. and H.F. Noller (1989) Interaction of tRNA with 23S RNA in the ribosomal A,P and E sites. *Cell* **57**, 585-597
12. Noller, H.F. (1993) Transfer RNA-Ribosomal RNA Interactions and Peptidyl Transferase. *FASEB J.* **7**, 87-89
13. Vester, B. and R.A. Garrett (1988) The importance of highly conserved nucleotides in the binding of chloramphenicol at the peptidyl transfer centre of Escherichia coli 23S ribosomal RNA. *EMBO J.* **7**, 3577-3587
14. Olsen, G.J., Woese, C.R. (1993) Ribosomal RNA - A Key to Phylogeny. *FASEB J.* **7**,113-123
15. Somerville, C.C., Jouannic, S., Martin, W.F., Kloareg, B., and Loiseaux-de Goer, S. (1993) Secondary structure and phylogeny of the chloroplast 23S rRNA gene from the brown alga *Pylaiella littoralis. Plant Mol. Biol.* **21**, 779-787
16. Vandepeer, Y., Neefs, J.M., Derijk, P., Dewachter, R. (1993) Evolution of Eukaryotes as Deduced from Small Ribosomal Subunit RNA Sequences. *Biochemical Systematics and Ecol.* **21**, 43-55
17. Escalante, A.A. Ayala, F.J. (1994) Phylogeny of the Malarial Genus Plasmodium, Derived from Ribosomal RNA Gene Sequences. Proc. Natl. Acad. Sci. USA 91,11373-11377
18. Kranz, H.D., Miks, D., Siegler, M.L., Capesius, I., Sensen, C.W. Huss, V.A.R. (1995) The Origin of Land Plants - Phylogenetic Relationships Among Charophytes, Bryophytes, and Vascular Plants Inferred from Complete Small Subunit Ribosomal RNA Gene Sequences. *J. Mol. Evol.* **41**, 74-84
19. Ludwig, W., Schleifer, K.H. (1994) Bacterial Phylogeny Based on 16S and 23S Ribosomal RNA Sequence Analysis. *Microbiol. Rev.* **15**, 155-173
20. Mcintyre, C.L., Winberg, B., Houchins, K., Appels, R., Baum, B.R. (1992) Relationships Between *Oryza* Species (Poaceae) Based on 5S DNA Sequences. *Plant Systematics and Evolution.* **183**, 249-264
21. Verma, A., Rattan, A., Tyagi, J.S. (1994) Development of a 23S Ribosomal RNA Based PCR Assay for the Detection of Mycobacteria. *Indian J. Biochemistry & Biophys.* **31**, 288-294
22. Takaiwa, F., Oono, K., Iida, Y., and M. Sugiura (1985) The complete nucleotide sequence of a rice 25S RNA gene. *Gene* **37**, 255-259

23. Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.R., Meng, BB.Y., Li, Y,Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome : Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217**, 185-194
24. Attasart, P., Atichartpongkul, S., Silawisas, K., and B. Sonthayanon. (1995) Ribosomal RNA Genes of *indica* rice mitochondria. In *Proceedings of the 11th FAOBMB Symposium [IUBMB Symposium no. 239]* (Svasti,J., Rimpanichayakit,V., Soontharos,S., Tassanakajorn,A., Limpaseni,T., Pongswasdi,P., Wilairat,P., Sonthayanon,B., Boonyawat,J., Packdibamrung,K., Kamolsiripichaiporn,S., eds.), Samakkhisan (Dokya) Pcl., Bangkok, pp. 120-125.
25. Maniatis, T., Fritsch, E.F., Sambrook, J. (1989) Molecular Cloning, Laboratory Manual. New York, Cold Spring Harbor Laboratory.
26. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and K. Struhl. (1987) Current protocols in molecular biology, Vol. 1, Green Publishing & Associates and Wiley Interscience, New York.
27. Feinberg, A.P. Vogelstein, B. (1984) A technique for radiolabelling DNA restriction fragments to a high specific activity. *Anal. Biochem.* **132**, 6-13
28. Kaiser,K., and Murray, N.E. (1985) The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. In *DNA Cloning : A practical approach* , Vol. 1 (Glover, D.M., ed.), IRL Press, Oxford, pp.1-47
29. Sanger, F. Nikeln,S. Coulson. A.R. (1977) DNA sequencing with chain-termination inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467
30. Hsiao KC. (1991) A fast and simple procedure for sequencing double stranded DNA with Sequenase. *Nucleic Acids Res.* **19(10)**, 2787.
31. Higgins, D.G., and Sharp, P.M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS* **5**, 151-153
32. Higgins, D.G., Bleasly, A., and Fuchs, R. (1991) Clustal V : improved software for multiple alignment. *CABIOS* **8**,189-191
33. Burks, C., Cassidy, M., Cinkosky, M.J., Cumella, K.E., Gilna, P, Hayden, J.E.D., Keen, G.M., Kelley, T.A., Kelly, M., Kristofferson, D., and Ryals, J. (1991) GenBank. *Nucleic Acids Res.* **19**, 2221-2225.
34. Brosius, J., Dull, T.J., Noller, H.F. (1980) Complete nucleotide sequence of a 23S ribosomal RNA gene from *Escherichia coli.* Proc. Natl. Acad. Sci. USA 77, 201-204
35. Edwards, K. and Kossel, H. (1981) The rRNA operon from Zea mays chloroplasts : nucleotide sequence of 23S rDNA and its homology with *E. coli* 23S rDNA. *Nucleic Acids Res.* **9**, 2853-2869
36. Zhang, Q., Saghai-Marof M.A., Lu, T.Y., and B.Z. Shen (1992) Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. *Theor. Appl. Genet.* **83**, 495-499
37. Ranjekar, P.K., Pethe, V.V., Chitnis, P.K., Lagu, M.D., Gupta, V.S. (1993) Use of RFLP in Detection of Genetic Diversity in Some *Indica* Rice Cultivars. *J. Plant Biochem. and Biotech.* **2**, 87-89
38. Fukuoka, S., Hosaka, K., and O. Kamijima (1992) Use of random amplified polymorphic DNAs (RAPDs) for identification of rice accession. *Jpn. J. Genet.* **67**, 243-252

บทคัดย่อ

รหัส cDNA ของยีน *rrn 23* ของข้าวอินดิก้าสายพันธุ์ขาวดอกมะลิ 105 ได้รับการศึกษา รหัสความยาว 1130 นิวคลีโอไทด์ที่หาได้และได้ตรวจสอบซ้ำแล้วในโคลน cDNA จำนวน 3 โคลน ครอบคลุมรหัสของยีนนี้ที่ตำแหน่งประมาณนิวคลีโอไทด์ที่ 1760 ถึง 2887 ซึ่งตรงกับโดเมนที่ 4 และ 5 ของ rRNA ชนิด 23S ของข้าว รหัสที่วิเคราะห์ได้นี้ พบว่ามีระดับความเหมือนกับยีนของข้าวจาโปนิก้า คิดเป็น 99.4% และกับยีนของข้าวโพด 99.0% ความแตกต่างของรหัสที่พบมีปรากฏในสองโดเมนของ rRNA นี้ ซึ่งถือกันว่าเป็นบริเวณที่ไม่มีโครงสร้างทุติยภูมิที่เห็นได้เด่นชัด รหัสของยีนที่ทราบจากการศึกษานี้ เชื่อว่าจะเป็นประโยชน์ในการจำแนกสายพันธุ์ข้าวได้ดี