A new rice repetitive DNA shows sequence homology to both 5S RNA and tRNA

Tiyun Wu and Ray Wu

Department of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853, USA

Received June 22, 1987; Accepted July 8, 1987

Accession no. Y00391

ABSTRACT

Moderately repetitive DNA sequences are found in the genomes of all eucaryotes that have been examined. We now report the discovery of a novel, transcribed, moderately repetitive DNA sequence in a higher plant which is different from any of the known repetitive DNA sequences from any organism. We isolated a rice cDNA clone which hybridizes to multiple bands on genomic blot analysis. The sequence of this 352 bp cDNA contains four regions of homology to the wheat phenylalanine tRNA, including the polymerase III-type promoter. Unexpectedly, two regions of the same 352 bp sequence also show homology to the wheat 5S RNA sequence. Using the cDNA as a probe, we have isolated six genomic clones which contain long tandem repeats of 355 bp sequence, and have sequenced nine repeat units. Our findings suggest that the rice repetitive sequence may be an amplified pseudogene with sequence homology to both 5S RNA and tRNA, but organized as long tandem repeats resembling 5S RNA genes. This is the first example showing homology between the sequences of a moderately repetitive DNA with unknown function and 5S RNA.

INTRODUCTION

A substantial fraction of the genome of higher eukaryotes contains arrays of moderately and highly repetitive sequences¹⁻⁹, the functions of which are unknown. Three general types of repetitive sequences can be distinguished (see refs. 1-5 and 9 for review): (a) short interspersed DNA sequences (SINES), including the human Alu family, which contain RNA polymerase III type promoters; (b) long interspersed repetitive DNA sequences (LINEs), usually 5-7 kb in length, which are interspersed with non-repetitive DNA sequences and are restricted to particular genomic regions; and (c) essentially identical short repeating units, called satellite DNAs, which are found associated with heterochromatic regions of chromosomes and are arranged in tandem arrays. There is no evidence that any of the latter two classes of repeated sequences are transcribed⁹.

The origin and evolution of these repetitive sequences also remain mostly unknown. The human Alu family was shown to be homologous to the 7SL RNA and was believed to have been derived from the 7SL RNA (or DNA)¹⁰. It was also reported that some mammalian repetitive sequence families, including the rodent type-II Alu family, have remarkable sequence homology with specific tRNAs⁸. Repetitive DNA sequences have also been studied in plants. One recent report showed that small tandemly repeated DNA sequences of higher

plants may have originated from a tRNA gene¹¹. Taken together, these results suggest that many highly repetitive sequences of higher eukaryotes may have evolved from specific tRNA genes and other class III genes^{4,5}. The intragenic RNA polymerase III promoter may initiate transcription of the repeated sequences, which has been proposed to facilitate the amplification by the retroposition mechanism⁴. Consistent with this hypothesis is that members of the human Alu and rodent type-II Alu families can be transcribed by RNA polymerase III in vivo and/or in vitro¹². However, it has not been demonstrated that any repetitive sequence families in higher plants are transcribed. In this study, we directed our effort to the identification of repetitive and transcribed sequences of a higher plant. Our results showed that a moderately repeated and transcribed sequence of the rice Oryza sativa has homology with both tRNA and 5S RNA.

MATERIALS AND METHODS

Isolation of total rice DNA

Five varieties of rice were used. Oryza sativa L., var. Labelle was from C.N. Bollich, Oryza sativa var. IR36, var. IR26 and Oryza australiensis were from T.T. Chang and R. Coffman, Oryza sativa var. Calrose 76 was from N. Rutger. IR36 is an indica line, and Calrose 76 is a japonica line¹³. Oryza sativa has the AA genome and Oryza australiensis has the EE genome¹⁴. Total DNA was isolated from different varieties of rice as described previously¹⁵.

Isolation of total rice RNA

Seeds of rice (O. sativa L. var. Labelle) were sterilized overnight with a 1:20 dilution of Clorox, rinsed thoroughly with sterile water, germinated, and grown in sterile water in the dark at 30°C for 14 days. The seedlings were cut into small pieces and homogenized in a Waring blender with liquid nitrogen until the tissue became a fine powder. A mortar and pestle were used to grind the powder in 10 ml phenol saturated with 0.1 M NaOAc (pH 5.2) and 10 ml TES buffer (50 mM Tris-HCl, pH 8.5, 10 mM EDTA, pH 8.0, 2% SDS) for 10 min. Phenol (10 ml) and 10 ml TES were added again, and the mixture ground for an additional 10 min. This step was repeated two more times. The homogenate was transferred into a centrifuge bottle and spun at 1,610 x g for 10 min. The supernatant was extracted twice with phenol, the sample mixed with 1/10 volume of 3 M NaOAc, pH 5.2, and the nucleic acids precipitated with two volumes of ethanol. The pellet was dissolved in 5 ml TE buffer, and 1.2 ml of 1 M NaCl was added to a final concentration of 0.2 M. The sample was centrifuged at 10,000 x g for 10 min, and the RNA was precipitated with two volumes of ethanol. The pellet was resuspended in 4 ml of water, and the sample was treated with 2 units of RQ₁ DNase in the presence of 40 units of RNasin at 37°C for 2 hrs. After extraction twice with phenol, the RNA was precipitated with ethanol and then dissolved in TE buffer. All

solutions used for RNA preparations were pretreated with 0.1% diethylpyrocarbonate overnight, and the glassware was baked at 200°C overnight.

Construction of rice genomic library and cDNA library

For the construction of the genomic library, total rice DNA (*O. sativa* var. IR26) was partially digested with *Sau*3A and ligated to *Bam*HI cut λ EMBL4 DNA¹⁶.

For the construction of the cDNA library, poly(A) RNA was isolated¹⁷ and double-stranded cDNA was synthesized¹⁸. The double-stranded cDNA was ligated to *Eco*RI linkers and then into *Eco*RI site of plasmid pUC19 as described by Wu *et. al.*¹⁹.

Restriction enzyme analyses and genomic blot hybridization

Restriction enzymes were purchased from New England Biolabs and Bethesda Research Laboratories. Digested DNA samples were fractionated by electrophoresis in 1% or 0.5% agarose gels. DNA fragments were transferred to a NYTRAN filter which was hybridized to ³²P-labeled RC48 DNA at 42°C and autoradiographed. The hybridization solution contained 50% formamide, 5X SSC, 5X Denhardt's solution, 1% SDS, and 100 μ g/ml denatured salmon sperm DNA.

Northern blot analysis of RNA transcripts

Total rice RNA was electrophoresed under denaturing conditions on 1% agarose-formaldehyde gels and then transferred to nitrocellulose filters. Hybridization conditions were the same as in the Southern blot hybridization, but the solution contained in addition 0.1% SDS.

DNA sequence analysis

DNA sequencing was performed by using dideoxynucleotide chain termination procedure adapted to single-stranded M13 phage DNA.

RESULTS

Identification of transcribed rice repetitive DNA

We are interested in studying repetitive DNA sequences which are transcribed in rice. We screened a cDNA library using ³²P-labeled total genomic rice DNA as the probe, and found several strongly-hybridizing clones (presumably representing repetitive DNAs); each gave a different pattern of hybridization to rice genomic DNA. One of them, RC48, was studied in detail. Figure 1A shows the pattern of genomic blot hybridization using labeled RC48 DNA as the probe and four varieties of rice (*Oryza sativa* L. var. Labelle, var. IR36 and var. Calrose 76, and *O. australiensis*), each digested with four different restriction enzymes. The hybridization patterns of three of the rice varieties are similar but not identical. For example, complete digestion of varieties with *Pst*I showed ladders of 355 bp repeats up to 20 units (7,000 bp) in size with all the intermediary sized units visible; the trimer and tetramer bands are the darkest. The DNA samples were completely digested with the restriction enzymes, and the observed ladder of repeats is the result of mutations or methylation which have altered the restriction sites. Digestion with *Hind*III gave fragments mainly above 5 kb in

Fig. 1. Genomic Southern analyses of different varieties of rice. (A) Total rice DNA (20 µg/lane) was digested with restriction enzymes as indicated (A., *AluI*; E., *EcoRI*; H., *HindIII*; P., *PstI*) and fractionated electrophoretically on a 1% agarose gel. All DNA samples (lanes 1-16) showed approximately the same intensity after staining the gel with ethidium bromide and visualizing under UV light (data not shown). *HindIII* and *EcoRI* digested λ DNA were run as size markers. The DNA samples were transferred to a NYTRAN filter for hybridization using 32 P labeled RC48 DNA as the probe. Lanes 1-4, DNA from *Oryza sativa* L. var. Labelle; lanes 5-8, DNA from *O. sativa* L. var. IR36; lanes 9-12, DNA from *O. australiensis*; and lanes 13-16, DNA from *O. sativa* var. Calrose 76. (B) Total rice genomic DNA (30 µg/lane) was digested with restriction enzymes as indicated (B, *BglII*; H, *HindIII*) and fractionated electrophoretically on a 0.5% agarose gel. T4 DNA, λ DNA, *EcoRI* and *HindIII* digested λ DNA were run as size markers. Samples 1-5 on the NYTRAN filter were probed with RC48 DNA and the hybridization conditions were the same as in Fig. 1A. Lane 1, undigested DNA. Lanes 2 and 3, DNA digested with 30 units and 90 units of *BglII*, respectively. Lanes 4 and 5, DNA digested with 30 units and 90 units of *HindIII*, respectively. Lanes 6, 7, 8 and 9, the RC48 probe was removed from the filter by boiling for 20 min in 0.01 x SSPE, 0.5% SDS, then the filter was hybridized with the 32 P-nick-translated rice actin-1 gene. Lane 6 corresponds to lane 2, lane 7 to lane 3, etc.

Labelle (Lane 3), but dimers to pentamers of 352 bp repeats were present in IR36 (lane 7). There was no hybridization between the RC48 probe and DNA from *O. australiensis*. Complete digestion of rice DNA was shown by the fact that adding three times more enzyme gave the same pattern. Moreover, adding λ phage DNA to the rice sample showed complete digestion of the λ DNA (data not presented).

To confirm the finding that this rice repetitive DNA is indeed transcribed, we carried out RNA hybridization experiments (Northern blot analysis). Using RC48 as a probe, we found

Fig. 2. RNA hybridization analysis (Northern blot). Total rice RNA (25 μ g/lane) was extracted from 14 day-old rice seedlings, fractionated electrophoretically on a 1% formaldehyde agarose gel, and then transferred to a nitrocellulose filter. **Lane 1**, hybridized with 32 P-nick-translated RC48 and autoradiographed for 2 hours. **Lane 2**, hybridized with 32 P-nick-translated rice actin-1 gene and autoradiographed for 20 hours.

at least four distinct hybridizing bands with sizes ranging from 1 to 4 kb as shown in Figure 2. The intensity of the four bands (Lane 1) together was at least five times higher than that of actin-1 mRNA (Lane 2). After adjusting for the differences in exposure time, the amount of RC48 mRNA appeared to be at least 50 times higher than that of rice actin-1 mRNA.

Sequence of the rice cDNA (RC48) and sequence comparison to 5S RNA and tRNA

To examine the sequence and structure of RC48 more closely, we determined its sequence (Figure 3A) and compared it with other known sequences. Two interesting points

Fig. 3. Comparison of the sequences of rice RC48 with wheat 5S and wheat tRNA^{Phe}. (A) The sequences of the wheat 5S gene are shown above the RC48 sequence, with copy I between positions 1-124 and copy II between positions 182 and 302 (shown in both part A and B). The locations of four copies of the wheat tRNA^{Phe} are shown below the RC48 sequence. Sequence homology with the rice RC48 is indicated with short vertical lines. Locations of the polymerase III-like promoter sequences (box A)²³ are shown. The sequences of wheat 5S RNA and wheat tRNA^{Phe} are taken from review articles^{20,24}. (B) Comparison of the rice RC48 sequence with two copies of wheat 5S RNA or four copies of wheat tRNA^{Phe}.

were revealed: (1) Two segments of RC48 (positions 1-124, 182-302) showed 45% and 42% sequence homology, respectively, with wheat 5S RNA²⁰. (2) Four segments of RC48 showed 41%, 45%, 53% and 43% sequence homology, respectively (Figure 3B), with wheat tRNA^{Phe}. Moreover, the RC48 includes the polymerase III-type promoter sequence²¹⁻²³

```

RC48 : 10      20      30      40      50      60
RG 1 : ----- A----- G----- G----- G-----
RG 2 : ----- A----- G----- G----- G-----
RG 5 : ----- A----- G----- G----- G-----
RG 6 : ----- A----- G----- G----- G-----
RG 7 : ----- A----- G----- G----- G-----
RG 8 : ----- A----- G----- G----- G-----
RG 9 : ----- A----- G----- G----- G-----
RG11A : ----- A----- G----- G----- G-----
RG11B : ----- A----- G----- G----- G-----
CON : 10      20      30      40      50      60
      GGTGTTTGTG AAGGCTGCA GAATCAGAG CCTCACCCTT GAGAGACTGG ATTTGGAGAG

RC48 : 70      80      90      100     110     120
RG 1 : ----- A----- C----- C----- G-----
RG 2 : ----- A----- C----- C----- G-----
RG 5 : ----- A----- C----- C----- G-----
RG 6 : ----- A----- C----- C----- G-----
RG 7 : ----- A----- C----- C----- G-----
RG 8 : ----- A----- C----- C----- G-----
RG 9 : ----- A----- C----- C----- G-----
RG11A : ----- A----- C----- C----- G-----
RG11B : ----- A----- C----- C----- G-----
CON : 70      80      90      100     110     120
      TTGGACCCCG AGAGATTTT CCACAAATTA CTTACCATAG GCCACGAAGC AGCGTTTGT

RC48 : 130     140     150     160     170     180
RG 1 : ----- A----- T----- C----- T-----
RG 2 : ----- A----- T----- C----- T-----
RG 5 : ----- A----- T----- C----- T-----
RG 6 : ----- A----- T----- C----- T-----
RG 7 : ----- A----- T----- C----- T-----
RG 8 : ----- A----- T----- C----- T-----
RG 9 : ----- A----- T----- C----- T-----
RG11A : ----- A----- T----- C----- T-----
RG11B : ----- A----- T----- C----- T-----
CON : 130     140     150     160     170     180
      TCAGCAGACG CTCGAATATC TGAGGCTCCA CAACTGGTTT TCAACBAGGT GCTTAGTATA

RC48 : 190     200     210     220     230     240
RG 1 : ----- T----- A----- G----- T-----
RG 2 : ----- T----- A----- G----- T-----
RG 5 : ----- T----- A----- G----- T-----
RG 6 : ----- T----- A----- G----- T-----
RG 7 : ----- T----- A----- G----- T-----
RG 8 : ----- T----- A----- G----- T-----
RG 9 : ----- T----- A----- G----- T-----
RG11A : ----- T----- A----- G----- T-----
RG11B : ----- T----- A----- G----- T-----
CON : 190     200     210     220     230     240
      GGGATGGCAC AGAGGACGTT AGGCTGGATA GCTTAGACGA TGTCAGCTGC AA*CG*GT+C

RC48 : 250     260     270     280     290     300
RG 1 : ----- T----- G----- A----- G----- A-----
RG 2 : ----- T----- G----- A----- G----- A-----
RG 5 : ----- T----- G----- A----- G----- A-----
RG 6 : ----- T----- G----- A----- G----- A-----
RG 7 : ----- T----- G----- A----- G----- A-----
RG 8 : ----- T----- G----- A----- G----- A-----
RG 9 : ----- T----- G----- A----- G----- A-----
RG11A : ----- T----- G----- A----- G----- A-----
RG11B : ----- T----- G----- A----- G----- A-----
CON : 250     260     270     280     290     300
      GACTAACCGC GCCGTTAAGT TTGGCTGGCG GGCBCGACAC TGTCGAATGC AGATGGAGAA

RC48 : 310     320     330     340     350
RG 1 : ----- C----- T----- G----- C----- C-----
RG 2 : ----- C----- T----- G----- C----- C-----
RG 5 : ----- C----- T----- G----- C----- C-----
RG 6 : ----- C----- T----- G----- C----- C-----
RG 7 : ----- C----- T----- G----- C----- C-----
RG 8 : ----- C----- T----- G----- C----- C-----
RG 9 : ----- C----- T----- G----- C----- C-----
RG11A : ----- C----- T----- G----- C----- C-----
RG11B : ----- C----- T----- G----- C----- C-----
CON : 310     320     330     340     350
      TTCTGTGCAA ATTAAACAT TGATAACTTT GCCTTCCAGT CATGGAATGC AGCTA

```

Fig. 4. Sequences of the cDNA clone RC48 and genomic clones RG1 to RG11. Only the sequence of RC48 and consensus (CON) is given. For the genomic clones, only those nucleotides which differ from the RC48 sequence are shown, and the identical sequences are represented by dashes. Asterisks indicate positions where a single base pair deletion occurred. Note that the sequence of some clones is not complete. RC1-RC9 are monomers of the repeat unit. RC11A and RC11B represent two monomer sequences within a dimer.

present in the tRNA gene²⁴. For example, the promoter box A sequences of copy 2, 3 and 4 all have 7 nucleotides out of 11 identical to that of the tRNA^{Phe}. Thus, RC48 may be a pseudogene related to two different types of sequences. tRNA-like sequences are frequently found in repetitive DNAs from other organisms, especially mammals^{4,5,8}. However, 5S RNA-like sequences have not yet been reported among repetitive DNA molecules. The

possibility that the rice repetitive DNA may be related to 5S RNA is strengthened by the following observations: (a) the overall sequence homology (e.g., nucleotides 1-124) is statistically highly significant (chi square = 39, $p < 0.0001$, based on 56 matches and 68 mismatches or unmatches); (b) in one region (positions 19-30), the homology between the rice RC48 and the wheat 5S RNA is 9 out of 12 nucleotides; and (c) rice RC48 includes a sequence, AAATTAAAA, identical to a part of the A/T rich transcriptional terminator sequence of the wheat 5S RNA gene⁹. This sequence is found between nucleotides 306 and 314 (marked with a heavy line above the rice RC48 sequence in Figure 3A) in a location expected for the terminator of copy II of the wheat 5S sequence.

In addition to DNA sequence homology to tRNA and 5S RNA, the rice repetitive DNA also resembles the 5S RNA gene in organization. Genomic blot analysis (on a 0.5% agarose gel) using total rice DNA digested with HindIII or BglII and RC48 as a probe showed that the length of over 50% of the rice repetitive sequences is around 160 kb (Figure 1B). Complete digestion of rice DNA with restriction enzymes was shown by the fact that when the filter was washed and probed again with rice actin-1 gene, the expected sizes of the actin gene-containing fragments were found (lanes 6-9), and no hybridization was found at the 160 kb region. Thus, the organization of this rice repetitive DNA in very long tandem repeats resembles that of the 5S RNA gene⁹, and suggests that it may be an amplified 5S RNA pseudogene.

Sequence of RC48-related genomic DNA

We next analyzed the genomic sequences related to the rice cDNA sequence. After screening about 20,000 plaques, we isolated six genomic clones from a λ EMBL4 rice genomic library using RC48 as a probe. DNAs from two clones were separately digested with EcoRI or SalI, and the fragments were cloned into M13 vectors. Seven monomer-sized fragments and one dimer-sized fragment were isolated and sequenced (Figure 4). The sequence homology between RC48 and any one of the genomic clones is between 90-91%. The homology between any two genomic clones varies between 91-96%. This result shows that all the genomic clones so far sequenced are closely related to one another and to the cDNA; however, we have not yet cloned the particular genomic repeat units from which the transcript in RC48 is derived. This is not surprising because quantitative Southern and dot blot assays indicate that there are over 1,000 copies of this 355 bp repeat unit per haploid rice genome.

Figure 5 shows the classical cloverleaf structure of the wheat phenylalanine tRNA²⁴ (W), and one of the four tRNA^{Phe}-like cloverleaf structures (copy 3) within the 355 bp repeat derived from the rice consensus sequence (R). The overall sequence homology is 53%. The general structure of the tRNA-like rice sequence including all the stem regions (AS, DS, ACS and TS), as well as the anticodon loop (ACL) and the pseudouridine loop (TL), resembles that of the wheat tRNA^{Phe}. The fact that the rice repetitive sequence can be folded into tRNA-like

Fig. 5. Comparison of the secondary structure of wheat phenylalanine tRNA²⁴ to copy 3 of tRNA-like structures in the consensus sequence of the rice repetitive DNA. (W) A typical cloverleaf structure of wheat tRNA^{Phe}. The 5' and 3' termini are labeled, as are the aminoacyl stem (AS), the dihydrouridine loop (DL), the anticodon loop (ACL), and the pseudouridine loop (TL). Uridine, dihydrouridine and pseudouridine residues are all written as U. (R) A tRNA-like structure derived from nucleotides #178-249 of the rice consensus sequence. Filled circles are drawn between bases in the cloverleaf structures to demonstrate regions of base-pairing. Open circles are drawn between guanine-uridine base pairs which are allowed in the analysis of secondary structure⁸. Bold-faced letters represent those nucleotides that are identical in W and R. Arrows in the DL region represent the invariant nucleotide in all polymerase III-type promoters²³.

cloverleaf structures with high level of resemblance to the secondary structure of wheat tRNA^{Phe} further strengthens the potential for evolutionary relationship. These results suggest that this rice repetitive sequence may be an amplified tRNA pseudogene.

DISCUSSION

The rice repetitive DNA shares certain characteristics of several types of plant repetitive DNAs. However, distinctive differences are found as follows. (1) Satellite DNAs with 177 bp repeat units were found in radish and several other Cruciferae^{11,25}. However, the satellite DNAs were not transcribed as shown by the lack of hybridization of total RNA to the 177 bp DNA probes²⁵. In contrast, the rice repetitive DNA is transcribed. (2) A highly repeated DNA sequence from *Arabidopsis thaliana* has been cloned and sequenced. The repeat units have an average length of 180 bp and are tandemly repeated in arrays larger than 50 kb²⁶. This sequence showed 44% homology with the rice repetitive DNA between nucleotides 1-180. The homology is lower between nucleotides 181-355. Whereas the rice repetitive sequence is transcribed, it is not known whether the *Arabidopsis* sequence is transcribed. (3) Our genomic blot experiment (Figure 1A) probed with the cDNA (RC48) gave patterns resembling

that of a highly repeated satellite DNA identified by Pental and Barnes²⁷. However, the latter DNA has not yet been further characterized, and thus it is not known if it is expressed or if it is related to the RC48 sequence. (4) 5S RNA genes of wheat are clustered in tandem arrays with repeating units 410 bp and 500 bp in length. Within a repeat unit, the sequence of a 120 bp 5S RNA gene and a 70 bp 5' proximal sequence are conserved between different repeat units, whereas the remainder of the noncoding spacer region varies in sequence⁹. In contrast, the rice repetitive DNA is arranged in clusters with repeating units 355 bp in length, with the entire 355 bp sequence showing high sequence homology among different repeating units.

Partial DNA sequence homology of the rice repetitive sequence to both tRNA and 5S RNA is surprising. In addition to sequence homology, resemblance of the potential secondary structure of the rice sequence to the secondary structure of tRNA on the one hand and to the organization of 5S RNA genes involving long tandem repeats on the other hand, raise the possibility that the rice repetitive DNA may be evolved from the common ancestor of 5S RNA and tRNA.

ACKNOWLEDGEMENTS

We thank Hua Xiao, Charles Aquadro, Prescott Deininger and Volker Vogt for helpful suggestions, Xiaoqing Silvia Wu for assistance in computer analysis, and Kimberly Reece for a clone of rice actin gene. This work was supported by research grants RF84066, Allocation No. 3, from the Rockefeller Foundation, and GM29179 from the National Institutes of Health.

REFERENCES

1. Schmid, C.W. and Jelinek, W.R. (1982) *Science* 216, 1065-1070.
2. Jelinek, W.R. and Schmid, C.W. (1982) *Ann. Rev. Biochem.* 51, 813-844.
3. Singer, M.F. (1982) *Int. Rev. Cytol.* 76, 67-112.
4. Rogers, J. (1985) *Int. Rev. Cytol.* 93, 187-279.
5. Weiner, A.M., Deininger, P.L. and Efstratiadis, A. (1986) *Ann. Rev. Biochem.* 55, 631-661.
6. Van Arsdell, S.W. *et al.* (1981) *Cell* 26, 11-17.
7. Jagadeeswaran, P., Forget, B.G. and Weissman, S.M. (1981) *Cell* 26, 141-142.
8. Daniels, G.R. and Deininger, P.L. (1985) *Nature* 317, 819-822.
9. Flavell, R. (1980) *Ann. Rev. Plant Physiol.* 31, 569-596.
10. Ullu, E. and Tschudi, C. (1984) *Nature* 312, 171-172.
11. Benslimane, A.A., Dron, M., Hartman, C. and Rode, A. (1986) *Nucleic Acids Res.* 14, 8111-8119.
12. Allan, M. and Paul, J. (1984) *Nucleic Acids Res.* 12, 1193-1200.
13. Tsunoda, S. (1984) in *Biology of Rice*, Tsunoda, S. and Takahashi, N. Eds., pp. 361-375, Elsevier, Amsterdam.
14. Morishima, H. (1984) in *Biology of Rice*, Tsunoda, S. and Takahashi, N. Eds., pp. 3-30, Elsevier, Amsterdam.
15. Moon, E., Kao, T.H. and Wu, R. (1987) *Nucleic Acids Res.* 15, 611-630.
16. Wu, R., Peng, Z., Kao, T., Moon, E. and Cai, Y. (1986) in *Rice Genetics*, Proc. International Rice Genetics Symposium, pp. 825-837, Manila, Philippines.
17. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory, New York.

18. Mansson, P.E., Hsu, D. and Stalker, D. (1985) *Mol. Gen. Genet.* 200, 356-361.
19. Wu, R., Wu, T.Y. and Ray, A. (1987) *Methods in Enzymol.* 152, 342-349.
20. Erdmann, V.A. (1982) *Nucleic Acids Res.* 10, r93-r115.
21. Daniels, G.R. *et al.* (1983) *Nucleic Acids Res.* 11, 7579-7593.
22. Daniels, G.R. and Deininger, P.L. (1985) *Nucleic Acids Res.* 11, 7595-7610.
23. Gabriella, G., Hofstetter, H. and Birnstiel, M.L. (1981) *Nature* 294, 626-631.
24. Gauss, D.H. and Sprinzl, M. (1983) *Nucleic Acids Res.* 11, r1-r53.
25. Grellet, F., Delcassa, D., Panabieres, F. and Delseny, M. (1986) *J. Mol. Biol.* 187, 495-507.
26. Martinez-Zapater, J.M., Estel and Somerville, C.R. (1986) *Mol. Gen. Genet.* 204, 417-423.
27. Pental, D. and Barnes, S.R. (1985) *Theor. Appl. Genet.* 70, 185-191.