

# Structure and gene expression of the mitochondrial genome of rice

M. Nakazono, N. Tsutsumi, and A. Hirai

We examined the structure and organization of the mitochondrial genome and the expression of the mitochondrial genes of rice (*Oryza sativa* cv Nipponbare). In particular, the entire set of the transferred chloroplast DNA sequences and a family of the palindromic repeated sequences that mainly exist in the intergenic spacer regions were identified. Furthermore, we determined the transcription initiation sites of rice mitochondrial genes by in vitro capping and ribonuclease protection assays and primer extension analysis. A consensus sequence was observed around the sites of initiation of transcription, indicating that the consensus motif is recognized as promoters of rice mitochondrial genes.

It has been known that the mitochondrial genome of higher plants is the largest and most complex among those of the eukaryotes. However, the plant mitochondrial genome does not contain many genes. Noncoding regions, the intergenic spacer regions, occupy a majority of the mitochondrial genome. There are a lot of pseudogenes derived from both mitochondrial and chloroplast genes, many small repeated sequences, and regulatory elements such as promoter and terminator elements in the intergenic spacer regions. In the present study, we investigated the structure and gene expression of the mitochondrial genome of rice, placing priority on the features of the intergenic spacer regions.

# Identification of the entire set of transferred chloroplast DNA sequences in the mitochondrial genome of rice (Nakazono and Hirai 1993)

## ctDNA sequences in the rice mitochondrial genome

The mitochondrial genomes of higher plants are much larger and more complex than those of other eukaryotic organisms (Newton 1988). It has been suggested that one of the reasons for such complexity might be the presence of many chloroplast sequences in the mtDNA of higher plants. Although there are many reports of chloroplast-like sequences in the mtDNA of higher plant, there are still few comprehensive surveys of these sequences. However, Stern and Palmer (1986) identified all locations of chloroplast-like sequences in the mtDNA of spinach by Southern hybridization.

We have carried out hybridization analysis using clone banks of chloroplast DNA (ctDNA) and mtDNA from rice (*Oryza sativa* cv Nipponbare) to examine the chloroplast sequences in the mtDNA of rice. Nucleotide sequences of hybridized fragments were determined and compared with the reported chloroplast sequences. We describe the entire set of chloroplast sequences in the rice mitochondrial genome with exact lengths of transferred fragments.

We found 16 chloroplast fragments in rice mtDNA that ranged from 32 bp to about 6.8 kb in length. As shown in Figure 1, the transfer of ctDNA occurred from

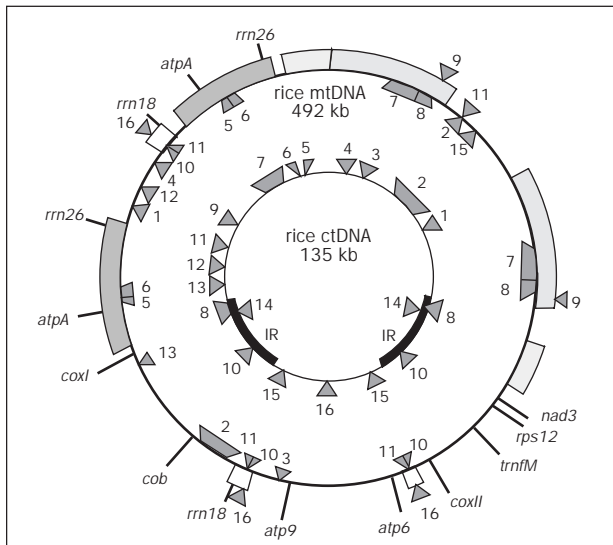


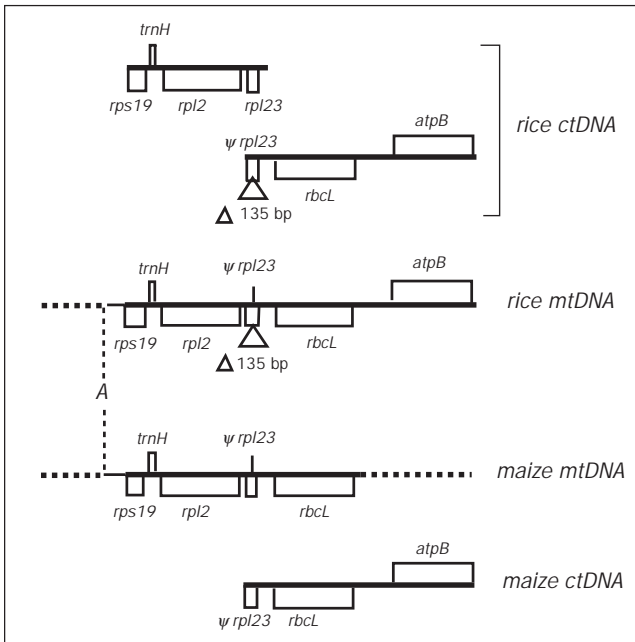
Fig. 1. The chloroplast DNA sequences in the rice mitochondrial genome. The inner circle indicates the rice chloroplast genome (Hirai et al 1985). Thick lines indicate the inverted repeat (IR) sequence. The outer circle is a master circle of the rice mitochondrial genome (Iwahashi et al 1992). The repeated sequences are indicated by boxes outside the master circle. The sequences homologous to ctDNA are indicated by shaded triangles with numbers. The numbers correspond to those in Table 1.



widely separated regions of ctDNA to widely separated sites in the mtDNA. Transferred chloroplast sequences are summarized in Table 1. The results show that about 6% (22 kb) of rice mtDNA, excluding repeated sequences, is made up of chloroplast sequences. Thus, about 19% of the rice ctDNA, omitting one of the inverted repeats, must have been transferred to the rice mtDNA. We did not find any common sequences or structures around the termini of chloroplast sequences that might explain the insertion of the ctDNA into the mtDNA.

### Rearranged ctDNA sequences in rice mtDNA

It was previously reported that the mitochondrial genome of rice (*O. sativa* cv Labelle) contains a rearranged cluster of chloroplast genes, namely, *rp12* plus *yrp123-rbcL-atpB-atpE-trnM-trnV* (Moon et al 1988). As shown in Figure 2, we also detected *rp12/yrp123-rbcL-atpB-atpE-trnM-trnV* in the mitochondrial genome of the cultivar Nipponbare. In the previous report, the terminus on the *rp12* side was not reported. However, this terminus was found to be located within *rps19* in our analysis. We compared rice mitochondrion-specific sequences at the border with available sequences



**Fig. 2.** Comparison between chloroplast sequences in the mitochondrial genomes of rice and maize. Thick lines indicate chloroplast sequences. Regions of high homology between rice and maize mitochondrion-specific sequences at the border of chloroplast sequences are shown by thin lines. The sequence upstream from position A in maize mitochondrial DNA was not available at the time of writing.

from maize (Iams et al 1985), and found 93% homology among sequences of 84 bases. The result indicates that the transfer of this region occurred before the divergence of rice and maize. By contrast, rice *yrp123* in ctDNA and mtDNA had a 135-bp deletion. A maize mitochondrial clone, pLSH20, containing *yrp123* was sequenced, and we found that neither maize gene had a deletion, as illustrated in Figure 2 (Bowman et al 1988). These facts reveal that this region was transferred from ctDNA after the divergence of rice and maize. Thus, it is suggested that *rps19-trnH-rp12-rp123* must have been transferred to the mtDNA before the divergence of rice and maize, and the region corresponding to *yrp123* must have been transferred after the divergence of rice and maize. Then the two regions must have become connected as a result of at least two recombination events, one in rice and the other in maize. This sequence must be a “hot spot” for recombination in mitochondria.

Insertion of palindromic repeated sequences (PRs) after divergence of genus *Oryza* from other Gramineae (Nakazono et al 1994)

### Number and location of PRs in the rice mitochondrial genome

When analyzing transferred sequences of chloroplast DNA in the mitochondrial genome of rice (*O. sativa* cv. Nipponbare), we found three small repeated sequences, which were about 60 bp in length, in the flanking regions of chloroplast-like (ct-like) sequences. To determine whether these repeats are widely distributed in the mitochondrial genome of rice, we performed plaque hybridization to lambda clones in a clone bank of rice mtDNA using a mixture of two oligonucleotides, which corresponded to sequences on the right (35 bases; IS-R, 5'-AAGAAAACGGATGCGCGTGCTAACGTTTCGCGCTA-3' and the left (36 bases; IS-L, 5'-CCAAACAAGCAACGGATTGAGCGCACTAGCGGAAA-3') sides of the small repeated sequence (Iwahashi et al 1992). It was found that rice mtDNA contained at least 10 copies of the small repeated sequence. We carried out subcloning and sequencing of the hybridizing regions. These small repeated sequences were distributed throughout the mitochondrial genome of rice (Nakazono et al 1994). It was predicted by computer-assisted analysis that the small repeated sequences had the potential ability to form stem-and-loop structures. Therefore, these repeats were designated the PRs (palindromic repeated sequences) of rice mtDNA (Fig. 3a). Sequencing data show that these repeats are from 60 to 66 bp in length and exhibit a high degree of homology to one another (Fig. 3b).

### A PR is located in the intron of the gene for ribosomal protein S3 (*rps3*) of rice mtDNA

We identified genes for ribosomal proteins S3 (*rps3*) and L16 (*rpl16*) in rice mtDNA. Both genes were not only closely linked but overlapped, as they do in maize mtDNA (Fig. 4; Hunt and Newton 1991, Nakazono et al 1995a). One of the PRs (361-PR; PRS included in clone No. 361) was located in the intron of *rps3* in rice mtDNA, but not in maize mtDNA (Fig. 4). It seemed likely that either a PR had been inserted in

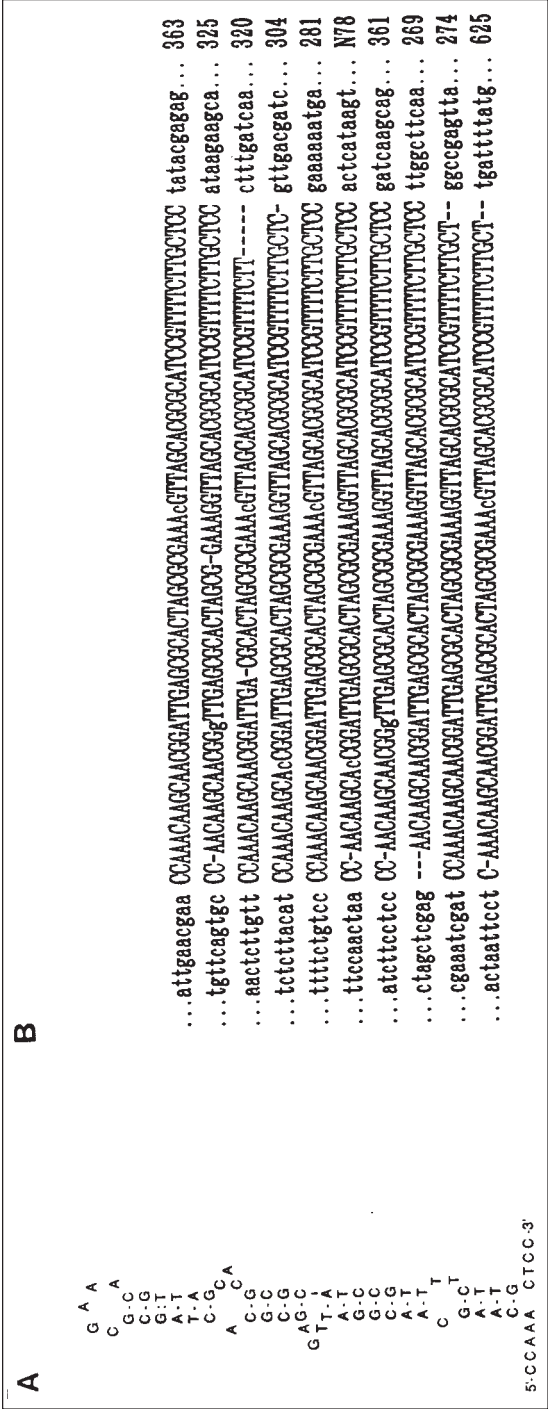


Fig. 3. a. Possible secondary structure of the small repeated sequence designated 363. The stem-and-loop structure was generated using GENETYX Software (SDC, Tokyo, Japan). b. Alignment of 10 small repeated sequences in the mitochondrial genome of rice. Capital letters indicate conserved nucleotide sequences.

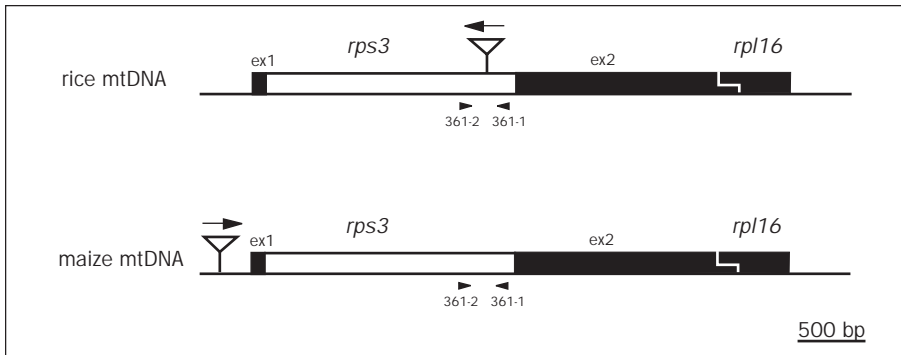


Fig. 4. Organization of *rps3* and *rpl16* in rice and maize mtDNAs. The closed boxes and the open boxes indicate the coding regions and the introns, respectively. ex1 and ex2 indicate exons 1 and 2 of *rps3*. The positions of each PRS are indicated by open triangles. Arrows indicate the directions of each PRS. Arrowheads indicate the positions of each primer (361-1, 361-2) used for amplification by PCR.

this region in rice or had been deleted from the corresponding location in maize. To examine these possibilities, we performed amplification by PCR using primers and total DNAs from various species of *Oryza* [*O. sativa* (Nipponbare), *O. rufipogon* (perennial type; W0120), *O. barthii* (W0822), *O. glaberrima* (W0025), *O. punctata* (W1515), *O. minuta* (W1319), *O. eichingeri* (W1521), *O. latifolia* (W0542), and *O. australiensis* (W0008)], as well as from maize, wheat, barley, and sorghum. The two primers [27 bases and 26 bases; designated 361-1 (5'-GAGGATCCTCTCTGAACCGTGCTAGAT-3') and 361-2 (5'-GCGGATCCAGAACAACCTCTGAACGTT-3'), respectively] corresponded to sequences in the two flanking regions of 361-PRS.

The anticipated 323-bp fragments were amplified in all of the species of *Oryza*, as shown in Figure 5a. However, the fragments amplified in the case of the other Gramineae were about 60 bases shorter than those of *Oryza* species. Southern hybridization was performed with the probes described above (IS-R and IS-L; Fig. 5b). Hybridization signals were obtained only from species of *Oryza*, showing that only fragments amplified from *Oryza* DNAs contained PRSs. We cloned the individual PCR-generated fragments of the other Gramineae in the pUC119 vector and determined their sequences. Only rice mtDNA had a PRS in this region, but the flanking sequences around the PRS were much the same among the five members of the Gramineae (Nakazono et al 1994). These results indicate that a PRS had been inserted in the intron of *rps3* of *Oryza* mtDNA after the divergence of *Oryza* from other Gramineae. Both ends of this PRS had incomplete direct repeats of nine bases (TCTT<sup>c</sup>CTCC). This evidence indicates that some PRSs may be capable of movement within the mitochondrial genomes of *Oryza*.

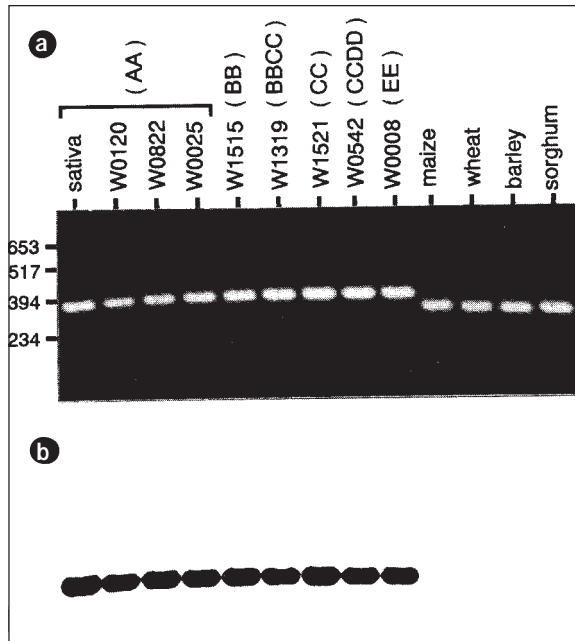
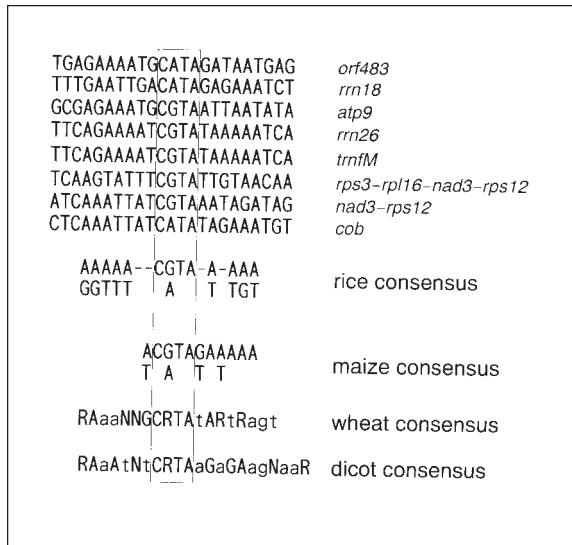


Fig. 5a and b. Ethidium bromide staining and Southern hybridization of PCR-generated fragments after electrophoresis. Amplifications by PCR were performed using primers 361-1 and 361-2. PCR-generated fragments were subjected to electrophoresis in a 1.2% agarose gel. AA, BB, BBCC, CC, CCDD, and EE indicate each nuclear genome-type of *Oryza*. The numbers on the left indicate sizes of fragments in basepairs (a). Southern hybridization was performed using a mixture of oligonucleotides IS-R and IS-L as probes (b).

## Identification of transcription initiation sites on the mitochondrial genome of rice (Nakazono et al 1995b)

The 5' end of a primary transcript, synthesized from the site of initiation of transcription in a plant organelle, can be specifically labeled *in vitro* by the capping enzyme guanylyltransferase (Auchincloss and Brown 1989, Vera and Sugiura 1992). The precise 5'-ends of primary transcripts of five genes (*orf483*, *rrn18*, *atp9*, *rrn26*, and *trnfM*) and one gene cluster (*rps3-rp116-nad3-rps12*) were investigated using *in vitro* capping/ribonuclease protection experiments and primer extension analysis (Nakazono et al 1995a,b,c). The transcription initiation site of the *cob* gene has been established by Kaleikau et al (1992). As shown in Figure 6, we aligned sequences around the initiation sites of these monocistronic and polycistronic transcripts (*orf483*, *rrn18*, *atp9*, *rrn26*, *trnfM*, *rps3-rp116-nad3-rps12*, and *cob*). The alignment revealed a high



**Fig. 6.** Alignment of the nucleotide sequences around the transcription initiation sites of rice mitochondrial genes and a gene cluster (*orf483*, *rrn18*, *atp9*, *rrn26*, *trnM*, *rps3-rpl16-nad3-rps12*, and *cob*). The consensus motif of the promoters of rice mitochondrial genes was deduced from the alignment. Consensus promoter sequences of the mitochondrial genes of maize, wheat, and dicots are taken from the reports by Mulligan et al (1991), Covello and Gray (1991), and Binder and Brennicke (1993), respectively. The box indicates the CRTA motif that is highly conserved in the promoter regions of monocots and dicots.

degree of similarity among sites, suggesting that the conserved motif [(A/G) (A/G) (A/T) (A/T) (A/T) NNC (G/A) TAN (A/T) N (A/T) (A/G) (A/T)] is recognized as the promoter of rice mitochondrial genes. This consensus sequence of rice exhibits a high degree of homology to those of other gramineous plants, i.e., maize (Mulligan et al 1991) and wheat (Covello and Gray 1991). In particular, the sequence C (G/A) TA, known as the CRTA motif, was found to be highly conserved in the promoter of rice mitochondrial genes examined by the *in vitro* capping method, as it is in those of other plants [maize (Mulligan et al 1991), wheat (Covello and Gray 1991), soybean (Brown et al 1991), and *Oenothera* (Binder and Brennicke 1993)]. We propose that the CRTA motif is necessary for the active transcription of the mitochondrial genes of rice, as it appears to be the case in the other plants examined.

To determine the number of promoters in the mitochondrial genome of rice, rice mtDNA digested by *Xba*I, *Sal*I, and *Bam*HI, respectively, was allowed to hybridize with *in vitro* capped mtRNA and was then treated with ribonucleases T1 and A. Thirteen to 16 signals (*Xba*I, 13 signals; *Sal*I and *Bam*HI, 16 signals each) were detected

(Nakazono et al 1995b). If one transcription initiation site is present in each detected fragment, this observation indicates that the mitochondrial genome of rice includes at least 16 promoters. As described by Binder and Brennicke (1993), transcription initiation sites of primary transcripts that are rapidly processed and/or degraded may be not detectable with this method. However, it is possible that signals that originate from different promoters might be duplicated in restriction fragments of a single size. Furthermore, we cannot rule out the possibility that some primary transcripts, transcribed from promoters in repeated sequences, might hybridize to several fragments. The locations of the individual promoters on the mitochondrial genome of rice should now be investigated.

## Cited references

- Auchincloss AH, Brown GG. 1989. Soybean mitochondrial transcripts capped in vitro with guanylyltransferase. *Biochem. Cell Biol.* 67:315-319.
- Binder S, Brennicke A. 1993. Transcription initiation sites in mitochondria of *Oenothera berteriana*. *J. Biol. Chem.* 268:7849-7855.
- Bowman CM, Barker RF, Dyer TA. 1988. In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, probably conserved by nonreciprocal recombination. *Curr. Genet.* 14:127-136.
- Brown GG, Auchincloss AH, Covello PC, Gray MW, Menassa R, Singh M. 1991. Characterization of transcription initiation sites on the soybean mitochondrial genome allows identification of a transcription-associated sequence motif. *Mol. Gen. Genet.* 228:345-355.
- Covello PS, Gray MW. 1991. Sequence analysis of wheat mitochondrial transcripts capped in vitro: definitive identification of transcription initiation sites. *Curr. Genet.* 20:245-251.
- Hirai A, Ishibashi T, Morikami A, Iwatsuki N, Shinozaki K, Sugiura M. 1985. Rice chloroplast DNA: a physical map and the location of the genes for the large sub-unit of ribulose 1,5-biphosphate carboxylase, and the 32KD photosystem II reaction center protein. *Theor. Appl. Genet.* 70:117-122.
- Hiratsuka J, Shimada, H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun C, Meng B, Li Y, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, 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.
- Hunt MD, Newton KJ. 1991. The NCS3 mutation: genetic evidence for the expression of ribosomal protein genes in *Zea mays* mitochondria. *EMBO J.* 10:1045-1052.
- Iams KP, Heckman JE, Sinclair JH. 1985. Sequence of histidyl tRNA, present as a chloroplast insert in mtDNA of *Zea mays*. *Plant Mol. Biol.* 4:225-232.
- Iwahashi M, Nakazono M, Kanno A, Sugino K, Ishibashi T, Hirai A. 1992. Genetic and physical maps and a clone bank of mitochondrial DNA from rice. *Theor. Appl. Genet.* 84:275-279.
- Kaleikau EK, Andre CP, Walbot V. 1992. Structure and expression of the rice mitochondrial apocytochrome b gene (*cob-1*) and pseudogene (*cob-2*). *Curr. Genet.* 22:463-470.
- Moon E, Kao T-H, Wu R. 1988. Rice mitochondrial genome contains a rearranged chloroplast gene cluster. *Mol. Gen. Genet.* 213:247-253.
- Mulligan RM, Leon P, Walbot V. 1991. Transcriptional and post-transcriptional regulation of maize mitochondrial gene expression. *Mol. Cell Biol.* 11:533-543.

- Nakazono M, Hirai A. 1993. Identification of the entire set of transferred chloroplast DNA sequences in the mitochondrial genome of rice. *Mol. Gen. Genet.* 236:341-346.
- Nakazono M, Kanno A, Tsutsumi N, Hirai A. 1994. Palindromic repeated sequences (*PRSs*) in the mitochondrial genome of rice: evidence for their insertion after divergence of the genus *Oryza* from the other Gramineae. *Plant Mol. Biol.* 24:273-281.
- Nakazono M, Itadani H, Wakasugi T, Tsutsumi N, Sugiura M, Hirai A. 1995a. The *rps3-rp116-nad3-rps12* gene cluster in rice mitochondrial DNA is transcribed from alternative promoters. *Curr. Genet.* 27:184-189.
- Nakazono M, Tsutsumi N, Hirai A. 1995b. The CRTA motif is present in the promoters of mitochondrial genes of rice. *Plant Sci.* 105:227-234.
- Nakazono M, Tsutsumi N, Sugiura M, Hirai A. 1995c. A small repeated sequence contains the transcription initiation sites for both *trnfM* and *rnm26* in rice mitochondria. *Plant Mol. Biol.* 28:343-346.
- Newton KJ. 1988. Plant mitochondrial genomes: organization, expression, and variation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 39:503-532.
- Stern DB, Palmer JD. 1986. Tripartite mitochondrial genome of spinach: physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. *Nucleic Acids Res.* 14:5651-5666.
- Vera A, Sugiura M. 1992. Combination of in vitro capping and ribonuclease protection improves the detection of transcription start sites in chloroplasts. *Plant Mol. Biol.* 19:309-311.

## Notes

*Authors' address:* Laboratory of Radiation Genetics, Faculty of Agriculture, The University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113, Japan.

*Acknowledgments:* The authors thank Drs. M. Sugiura, T. Wakasugi, K. Kadowaki, and A. Kanno for helpful discussions. This research was supported by grants-in-aid from the Ministry of Education, Science and Culture in Japan.

*Citation:* [IRRI] International Rice Research Institute. 1996. Rice genetics III. Proceedings of the Third International Rice Genetics Symposium, 16-20 Oct 1995. Manila (Philippines): IRRI.