Precise characterization of rDNA genes by intraspecies and inter-loci comparison of rDNA sequences and biochemical analysis of ribosomal RNA molecules in Agrobacterium tumefaciens

J-ney Bautista-Zapanta¹, Kazuo Yoshida² and Katsunori Suzuki¹*¹

¹Department of Biological Science, Graduate School of Science, Hiroshima University, Higashi Hiroshima 739-8526, Japan
²Laboratory for Food & Molecular Microbiology, Yasuda Women’s University, Yasu Higashi 6-13-1, Asaminami-ku, Hiroshima, 731-0153, Japan

(Received 16 September 2004, accepted 28 January 2005)

Annotation of rRNA genes has been incomplete in Agrobacterium species although a number of Agrobacterial rDNA fragments have been sequenced. In this study, precise characterization of rRNA operons (rrn) was carried out in two biovar 1 strains, C58 and MAFF301001. Complete DNA sequencing of four rrns in MAFF301001 indicated that each operon codes for 16S, 23S and 5S rRNA as well as three tRNAs, trn¹lle, trn¹ala and trn¹met. The genes and 16S-23S ITS of a given locus were exactly identical with those in the other three loci, except for a T-base loss in the 23S rRNA gene of rrnA and in the 5S rRNA gene of rrnB. Comparison with the four C58 rDNAs available in the DNA database indicated extensive sequence and size variations in the 23S rRNA gene, suggesting the presence of an intervening sequence (IVS). Biochemical RNA analysis, including Northern hybridization and 5‘ end mapping, in MAFF301001 revealed 2886-base and 2571-base precursors, two 1.3-kb major fragments, a 150-base fragment and removal of an IVS for 23S rRNA. We confirmed similar biochemical characteristics in the C58 strain. The features of rDNA detected here enable correction of previously reported information about Agrobacterial rRNAs and rRNA genes and should be useful for phylogenetic considerations.

Key words: Agrobacteria, ITS, ribosome, RNA processing, rrn operon

INTRODUCTION

Agrobacterium species are members of the Rhizobiaceae family under the alpha subgroup of Proteobacteria. Agrobacterium species are well known as the causative agents of crown gall disease and/or hairy root disease, and are useful for applications to plant transformation.

Because of their importance in phytopathology, more than 100 ribosomal RNA (rRNA) gene sequences were determined for use in identifying and classifying pathogenic strains, and the resultant data are available in the databases. In spite of the accumulation of the nucleotide sequence data, the information is fragmentary since biochemical data are limited and have not been examined in detail and the sequence data were obtained by performing PCR amplification of the conserved parts of the rRNA gene. Complete sequencing of an operon was reported for a few strains of Agrobacterium species. In biovar 3 (A. vitis) strains S4 and NCPPB3554, only one rRNA operon (rrn) out of four has been sequenced and characterized (Otten et al., 1996). Recent studies by Wood et al. (2001) and Goodner et al. (2001) have provided nucleotide sequences of four rrn operons in A. tumefaciens C58. However, rrnA gene structures were only predicted using the DNA sequence data. Accurate annotation using experimental data including 5‘ end mapping is still lacking in this species. In order to establish a reliable structural basis not only for taxonomy and biochemical purposes but also for evolutionary studies, extensive characterization of the operons is essential in the species.

Bacterial rRNAs, namely 16S, 23S and 5S rRNAs, are encoded by the operons in general (King et al., 1986). rRNA gene sequences are most frequently used to characterize the evolutionary history of organisms (Hori and Osawa, 1979; Woese, 1987). rRNA gene sequences are a powerful tool for determining phylogenetic relationships among organisms at higher taxa as well as at the species
level (Weisburg et al., 1991; Yanagi and Yamasato, 1993). Internal transcribed spacer (ITS) in the rrn operon has also been given significant attention for use in classification down to the intra-species level because it shows much higher variation than the coding sequence regions in the operons. ITS regions contain one or more tRNA genes (Anton et al., 1998; Casanova et al., 2001).

In some bacteria, 23S rRNA is processed to form much shorter molecules. The processing is triggered by the presence of an intervening sequence (IVS) in 23S rRNA. IVS was originally identified in Salmonella species (Burgin et al., 1990), and then reported in several members in γ-Proteobacteria. Recently, it has been reported that IVS is also present in two α-Proteobacteria members: Brucella melitensis (Bricker, 2000) and Rhizobium species (Selenska-Pobell and Evguenieva-Hackenberg, 1995). The IVS segment is removed, and further fragmentation occurs during rRNA processing (Zahn et al., 1999; Zahn et al., 2000; Zahn et al., 2001). The diversity of rRNA and rDNA described above has prompted extensive characterization and comparison. However, only a little attention has been paid to the precise features of rRNAs and their genes in Agrobacteria.

In this study, we characterized four ribosomal RNA operons (rrn) in a biovar 1 strain A. tumefaciens MAFF-301001, whose Ti plasmid was completely sequenced previously (Suzuki et al., 2000), and compared the data with those of another biovar 1 strain, C58. The resulting intra-species comparison predicted the presence of the IVS. Biochemical examination detected fragmentation of 23S rRNA, localized the 5' ends of rRNA molecules, detected the IVS, and corrected previous annotations among Agrobacteria. A preliminary account of this study was reported previously (Bautista-Zapanta, et al., 2002).

**MATERIALS AND METHODS**

**Bacterial strains and plasmids.** Biovar 1 A. tumefaciens strain MAFF301001 was used throughout this study. Another biovar 1 strain, C58C1, and Escherichia coli strain HB101 were used in part for comparative reference. E. coli SURE was used as the host strain for cloning DNA. Plasmid pBluescript II KS+ was used for subcloning of target DNA fragments. Recombinant BAC clones containing rrn genes were screened from the genomic library of MAFF301001 strain (De Costa et al., 2001) by colony hybridization using the 16S rRNA gene as a probe. For stable maintenance of BAC recombinant clones in E. coli, 0.2% glucose was added to LB medium.

**Pulsed field gel electrophoresis (PFGE), genomic DNA preparation and Southern hybridization.** Preparation of total genomic DNA in agarose plugs, digestion of the embedded DNA with macrorestriction endonucleases Pmel and Sva1, and subsequent separation by PFGE were carried out as previously described (Suzuki et al., 2001). PFGE was performed using a Pulsaphore apparatus (Pharmacia) with 200 V and 70 sec pulse time for 24 hours at 14°C. Southern hybridization was carried out using α-32P-labeled probes prepared by the random primer labeling method.

**Oligonucleotide primers.** The following oligonucleotide primers were used to sequence the two strands of the four rrn operons: 16S869F (5'-GCCCGCAAGGGGAG3'-3'), 16S509F (5'-GGCGTAAAGCGACGC3'-3'), 16S929F (5'-GCTTTGAAATTCGGGG3'-3'), 16S1374F (5'-CCCGAATGTCGCGCC3'-3'), 16S448R (5'-GCTCGCAAGAATGGCACC3'-3'), 16S993R (5'-CGACAGCGCTGAGCACCC3'-3'), 16S1451R (5'-GGTCTCCCTACGGCTACC3'-3'), 16S1853F (5'-GTTCAAGTCTCTCCGGG3'-3'), 23S241F (5'-AGCGAAGCCAGGAACCTG3'-3'), 23S660F (5'-AGCACATCTAAGGGG3'-3'), 23S1136F (5'-AGGTCCTTCAAAGAAGG3'-3'), 23S1611F (5'-CCTAGGCGCTTGAG3AAG3'-3'), 23S2061F (5'-AGGTATGTTGCTTCGACT3'-3'), 23S2439F (5'-TGGGCACTGCTGATGCA3'-3'), 23S1208R (5'-TCCGTTGCGATGGCCTT3'-3'), 23S2437R (5'-TGCAGACTGAGGCTTCCG3'-3'), 23S1894R (5'-GGTCTACCCCTGAGCTGTC3'-3') and 23S283R (5'-GATTCTCCGTGGATGCTTCC3'-3'). These oligonucleotide primers were designed based on the conserved regions of 16S and 23S rDNA of Agrobacterium biovar 1 strains in the DNA databases. For nucleotide sequencing of the remaining 1-kbp region covering 5S rDNA, trn3det and their flanking regions, the following primers were designed successively during primer walking experiments: rrn5616F (5'-TTCTGTCTTCACGGGCC3'-3'), rrn3F (5'-GCTATAAATAATACGT3'-3'), 16S1136F (5'-AGGTCCTTCAAAGAAGG3'-3'), 23S1853F (5'-GTTCAAGTCTCTCCGGG3'-3'), 16S509F (5'-GCCCGCAAGGGGAG3'-3'), 23S241F (5'-AGCGAAGCCAGGAACCTG3'-3'), 23S1208R (5'-GCTCGCAAGAATGGCACC3'-3'). The following oligonucleotide primers were used to prepare DNA fragments for Northern hybridization probes: 23S241F and 23S1208R to detect a proximal 1.3-kb fragment of 23S rRNA, Maf23Sup150Fw (5'-GTTTTGAGGACGCTG3'-3') and MafIVS129Fw (5'-TGTTTTGGGAGGCGTG3'-3') to detect the 129-base putative IVS. For primer extension analyses, the following oligonucleotide primers were used: 16SMaf89R (5'-CGCGTTACTCGCTG3'-3'), 23S1894R (5'-GGTCTACCCCTGAGCTGTC3'-3') and 23S283R (5'-GATTCTCCGTGGATGCTTCC3'-3'). For primer walking experiments, the following oligonucleotide primers were used: 16SMaf89R (5'-CGCGTTACTCGCTG3'-3'), 23S1894R (5'-GGTCTACCCCTGAGCTGTC3'-3') and 23S283R (5'-GATTCTCCGTGGATGCTTCC3'-3'). For Northern hybridization, a 2.8-kbp HindIII fragment that contained the 16S rRNA gene and spacers was excised from a fosmid clone of each locus. Each fragment was subcloned in pBluescript II
KS" and used as a template for sequencing. The region corresponding to 23S rRNA gene was PCR-amplified using 16S1374F and 23S2437R primers and fosmid DNA harboring each locus. Templates used to cover 1-kbp regions downstream of the 23S rRNA gene and upstream of the 16S rRNA were prepared by inverse PCR amplification using a self-ligated BanIII fragment of each fosmid DNA. Prior to self-ligation, the specific BanIII fragment containing each intact rRNA operon of the four rrn loci was identified by Southern hybridization analysis. DNA sequencing was carried out using a Dynamic ET terminator sequencing kit (Amersham Bioscience, Piscataway, New Jersey) according to the manufacturer's instructions followed by analysis in an ABI Prism Model 310 automatic sequencer (Applied Biosystems, Foster City, California).

Computer analysis. Raw DNA sequencing data were edited using the software SeqEd version 3.0 (Applied Biosystems) and assembled into contigs with the software Sequencer version 3.0 (Gene Codes Corporation, Inc., Ann Arbor, Michigan). The Genetyx Mac program (Software Development, Tokyo) was used to confirm the presence of tRNAs and their cloverleaf structures. Comparative analyses of the nucleotide sequence data were carried out using the expanded ClustalW alignment program (DDBJ) and maximum matching with the DNAsis software (Hitachi Software Engineering Co., Japan).

Isolation of RNA, denaturing gel electrophoresis and Northern hybridization analyses. *A. tumefaciens* MAFF301001 and C58C1 were cultured in 40 mL of LB broth with shaking at 28°C until exponential phase. Cells were collected from the culture by centrifugation and ribosomal particles were extracted as recommended by Liveris et al. (1991). To extract rRNA, ribosome pellets were dissolved in acetate/SDS solution (1% SDS in 10 mM EDTA, 50 mM sodium acetate) and boiled for 5 minutes following the method described by Rivas et al. (2001). The extracted rRNA was treated with DNase I to remove contaminating DNA using standard methods. Purified samples were stored at −80°C until used. Purity and concentration of RNA were evaluated spectrophotometrically (Sambrook and Russell, 2001).

RNA solution containing 0.5 μg of rRNA was analysed by denaturation gel electrophoresis in 1.5% agarose gels containing 7% formaldehyde and MOPS buffer (20 mM MOPS, 2 mM sodium acetate, 1 mM EDTA (pH7.0) as described by Sambrook and Russell (2001). A commercially available RNA size marker (Toyobo Co., Osaka) was used to estimate the size of each rRNA molecule. Total RNA extracted from *E. coli* HB101 was used as a control.

Electrophoresed RNA was stained with ethidium bromide to visualize RNA bands and subsequently blotted onto a nylon membrane (Biodyne A, Pall City) by capillary transfer overnight using 20XSSC buffer (3 M sodium chloride, 0.3 M sodium citrate (pH 7.0)). RNA was fixed to the damp membrane under UV light, and the membrane was stored at −80°C until use. Northern blots were prehybridized with hybridization solution (0.5 M sodium phosphate (pH 7.2), 1 mM EDTA (pH 7.0), 7% SDS (w/v) and 39 μg/mL salmon sperm DNA) and prehybridized at 60°C before addition of probes.

Mapping 5′ end of 16S and 23S rRNA molecules. To map the 5′ ends of rRNA molecules, primer extension was performed using total RNA as template, an oligonucleotide primer and Omniscript reverse transcriptase (QIAGEN Incorporated, Valencia, California) following the manufacturer’s instruction. The same oligonucleotide was also used in parallel for the cycle sequencing reaction against an rrnA fragment to serve as a size marker. The sequencing reaction and reverse transcription reaction products were denatured and electrophoresed in an 8% polyacrylamide gel containing 8 M urea for 3 hours at 1900 V. The gel was transferred onto absorbent paper and dried in vacuo. Autoradiographic exposure to imaging plates and subsequent detection were carried out using the BAS2000 Imaging System (Fuji film, Tokyo).

Sequence data deposition. The complete nucleotide sequence and annotation data of the four rrn loci in MAFF301001 strain presented in this study were deposited in the DNA databases (DDBJ/EMBL/Genbank) with accession numbers AB102732 for rrnA, AB102733 for rrnB, AB102734 for rrnC and AB102735 for rrnD.

RESULTS

Identification of four rDNA loci in the MAFF301001 genome. Our previous study showed that three regions in the physical map of the total genome of *A. tumefaciens* MAFF301001 contain rrn loci (Suzuki et al., 2001), i.e., rrnA in Swal#3-Pmel#3 fragment in the linear chromosome, rrnB in Swal#1-Pmel#2 and rrnC in Swal5-Pmel#2 of the circular chromosome (Fig. 1A and B). We found 20 rrn-containing fosmid clones in the genomic library of the strain as described in Materials and Methods. We identified four groups among the 20 rrn-containing clones by BanIII and HinIII digestion of each fosmid clone DNAs followed by a search for common fragment profiles. Four representative clones, one for each group, namely, 2E4, 1C2, 2M13 and 1O21, were chosen and used to assign clones to specific rrn loci.

To assign each representative clone to a specific rrn locus, T7- and Sp6-end regions of an insert in each representative clone were hybridized with a PFGE Southern blot of Swal1- and Pmel-digested total genomic DNA. The 1C2 clone was assigned to rrnB in the circular chromo-
somal SwaI#1-Pmel#2 region, while 2M13 was assigned to \( \text{rrnC} \) in the circular chromosomal SwaI#5-Pmel#2 region. Fosmid clones 2E4 and 1021 gave positive signals in the linear chromosomal SwaI#3-Pmel#3 region. Contiguous alignment between fosmid- and phage-clones in this region (Fig. 1C) enabled us to find two different \( \text{rrn} \) loci, hence, a new locus, \( \text{rrnD} \), was assigned to this region. 2E4 and 1021 were assigned to \( \text{rrnA} \) and \( \text{rrnD} \), respectively, as shown in Fig. 1. Hence, strain MAFF301001 harbors two \( \text{rrn} \) loci in the linear chromosome and two loci in the circular chromosome.

Comparison of nucleotide sequences among the four rDNA operons in MAFF301001. The complete nucleotide sequence of each of the four \( \text{rrn} \) loci was determined using the four fosmid clones. The gene organization for each \( \text{rrn} \) was predicted from the sequence data as described in Materials and Methods. As shown in Fig. 2A, each operon harbored the sequences for 16S rRNA, 23S-5S ITS, 5S rRNA gene and \( \text{trn} \) genes. The overall organization of each \( \text{rrn} \) in \( A. \) \textit{tumefaciens} was similar to those of other \textit{Rhizobiaceae} strains.

The nucleotide sequences of the genes were identical among the four \( \text{rrn} \) operons throughout the 5611-bp sequence starting from 180 bp upstream of the 16S rRNA gene to the 23S rRNA gene of \( \text{rrnA} \) (illustrated as a white dot in Fig. 2A). There were no nucleotide sequence differences in the 16S-23S ITS among the four \( \text{rrns} \). A single base difference (loss of \( T \) in the 5S rRNA gene of \( \text{rrnB} \) (illustrated as a white dot in Fig. 2A) was detected, and a conversion (C→A) in the spacer region between the 5S rRNA gene and \( \text{trn}^{\text{Met}} \) of \( \text{rrnA} \) was present (illustrated as black dots in Fig. 2A). In contrast to the above perfectly conserved 16S-23S ITS, nucleotide sequence variation was observed in the 23S-5S ITS among the four loci.

Comparison of four rDNA operons in C58 strain. The organization of the \( \text{rrn} \) operons of the C58 strain shown in the database is identical with that of MAFF301001. Comparison of the four \( \text{rrns} \) in C58 carried out in this study revealed that the four \( \text{rrns} \) of C58 could be classified into two groups. Three \( \text{rrn} \) operons, “operon 1”, “operon 2” and “operon 3” were highly conserved in the region from 62 bp upstream of the 16S rRNA gene to \( \text{trn}^{\text{Met}} \). The remaining \( \text{rrn} \) “operon 4” was identical with the other three \( \text{rrn} \) operons in its 16S rRNA genes as well as \( \text{trn}^{\text{Ile}}, \text{trn}^{\text{Ala}} \) and \( \text{trn}^{\text{Met}} \). However, its 16S-23S ITS was shorter by 83 bp and its sequence differed compared with those of the other three operons (illustrated as black dots in Fig. 2B’). The 23S rRNA gene of “operon 4” also contained 13-bp and 9-bp losses and sequences unrelated to those of the other three operons (white dots in Fig. 2B’). The 23S-5S ITS showed sequences unrelated to those of the other three operons (black dots in Fig. 2B’). In the region upstream of the 16S rRNA gene, only the 62-bp sequence was highly conserved, and sequences further upstream differed among the four operons. Just downstream of \( \text{trn}^{\text{Met}} \), there was no sequence conservation among the four \( \text{rrn} \) loci in either strain (data not shown).

Comparison of rDNA operons between strains MAFF301001 and C58. The gene assignment data of two \( \text{rrn} \) loci, “operon 2” and “operon 4” of C58, were taken from the DNA database (accession nos. NC_003304 & NC_003305 for “operon 2” and “operon 4”, respectively) and compared with the four \( \text{rrns} \) of MAFF301001. The comparison indicated that “operon 4” of C58 was more similar to the four \( \text{rrns} \) of MAFF301001 than to “operon 2” of C58 (data not shown). In the 16S-23S ITS, the sequences differed among the three \( \text{rrns} \).

The nucleotide sequence data of \( \text{rrnD} \) of MAFF301001 and “operon 2” and “operon 4” of C58 were compared by maximum matching. As shown in Table 1, the predicted
genes and intervening sequences in Agrobacteria

... size and degree of similarity of each gene was lower than expected between the two strains when we employed the C58 annotation from the database. The gene size discrepancies between the two strains were resolved in most part by recruiting some extra nucleotide sequences from the ITS regions of C58, as shown in Table 1. When the genes were compared again using the corrected annotation data, 16S rRNA of MAFF301001 was 98% identical with that of C58. 23S rRNA of MAFF301001 was 95% and 96% identical with that of C58 “operon 2” and “operon 4” and MAFF301001 5S rRNA was 97% and 99% similar with that of C58 “operon 2” and “operon 4”, respectively. The ITS regions had the least identity between the two strains (ranging from 35% to 76%) as shown in Table 1. Interestingly, a 129-bp sequence of MAFF301001 that is located 150 bp downstream from the 5' end of MAFF301001 23S rRNA gene was not related to a C58 41-bp sequence that is upstream of the 23S rRNA gene for both “operon 2” and “operon 4” in the previous annotations, as shown in Fig. 3. We predicted this region to be an intervening sequence (IVS) in 23S rRNA. The putative IVS sequences and their sizes were highly variable between C58 and MAFF301001. In contrast, the IVS sequence was conserved among the four operons in each strain.

Homology search in the DNA database for the 129-bp IVS of MAFF301001 demonstrated the highest Z-score (639) with the 23S rRNA sequence data of A. rubi strain LMG156, though no IVS had been assigned yet to the A. rubi's data.

**Fragmentation of 23S rRNA molecule and removal of the putative IVS.** Pioneering work by Schuch and Loening (1975) indicated that there are two types of 23S rRNA, 0.52 × 10^6 and 0.48 × 10^6 daltons in size, in A. tumefaciens strain B6. In addition, their radiolabeling experiment in vivo showed a precursor with a molecular size similar to that of E. coli 23S rRNA. To confirm whether similar molecules are present in MAFF301001 cells, rRNA was extracted and separated by denaturing gel electrophoresis.

**Fig. 4A** shows a ladder of MAFF301001 RNA consisting of three visible bands (2.6 kb, 1.5 kb and 1.3 kb), which were also observed in the C58 RNA preparation in this study (data not shown). The 1.5-kb band represents 16S rRNA molecule, while the 2.6-kb molecule corresponds to 23S rRNA, as shown by Northern hybridization analysis using a probe consisting of a fragment of the 23S rRNA gene (Fig. 4A, right panel). The latter (2.6 kb) was...
slightly smaller than *E. coli* 23S rRNA (2.9 kb).

The 1.3-kb RNA band appeared to be an extra fragment. A Northern hybridization experiment (Fig. 4A, right panel) showed that the 1.3-kb extra RNA band was derived from 23S rRNA. Its size is half the size of the 2.6-kb molecule and its intensity was two-fold stronger than that of the 1.5-kb 16S rRNA band. These data suggest that the 1.3-kb fragment resulted from fragmentation of the 2.6-kb 23S rRNA, which, if cut at its center, would produce two 1.3-kb segments, namely, a 1.3-kb fragment proximal and another 1.3-kb fragment distal to the 5' end of the 2.6-kb molecule (Fig. 4C). As shown in

<table>
<thead>
<tr>
<th>Gene/ region</th>
<th>MAFF301001 Length (bp)</th>
<th>Operon 2 Length (bp)</th>
<th>Identity (%)</th>
<th>Operon 4 Length (bp)</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>1495</td>
<td>1492 (1495*)</td>
<td>98 (98')</td>
<td>1492 (1495*)</td>
<td>98 (98')</td>
</tr>
<tr>
<td>16S -23S ITS</td>
<td>1034</td>
<td>1487 (1289*)</td>
<td>58 (67')</td>
<td>1406 (1208*)</td>
<td>62 (71')</td>
</tr>
<tr>
<td>trn^{Ile}</td>
<td>77</td>
<td>80 (77*)</td>
<td>92 (96')</td>
<td>80 (77*)</td>
<td>92 (96')</td>
</tr>
<tr>
<td>trn^{Ala}</td>
<td>76</td>
<td>79 (76*)</td>
<td>96 (100')</td>
<td>79 (76*)</td>
<td>96 (100')</td>
</tr>
<tr>
<td>23S rRNA</td>
<td>2571(^a), 2886(^b)</td>
<td>2257 (2821*)</td>
<td>75 (95')</td>
<td>2235 (2799*)</td>
<td>76 (96')</td>
</tr>
<tr>
<td>IVS</td>
<td>129</td>
<td>- (41*)</td>
<td>- (31')</td>
<td>- (41*)</td>
<td>- (31')</td>
</tr>
<tr>
<td>23S-5S ITS</td>
<td>239, 244(^c)</td>
<td>630 (258*)</td>
<td>35 (76')</td>
<td>630 (258*)</td>
<td>35 (74')</td>
</tr>
<tr>
<td>5S rRNA</td>
<td>143</td>
<td>123 (143*)</td>
<td>84 (97')</td>
<td>123 (143*)</td>
<td>85 (98')</td>
</tr>
<tr>
<td>trn^{Met}</td>
<td>77</td>
<td>79 (77*)</td>
<td>95 (100')</td>
<td>79 (77*)</td>
<td>95 (100')</td>
</tr>
</tbody>
</table>

Nucleotide sequence identity was counted by comparison between *rrnD* of MAFF301001 (accession no. AB102735) and/or the "operon 2" and "operon 4" of C58 with accession nos. NC_003304 & NC_003305, respectively. \(^a\) Mature 23S rRNA molecule shown in Fig. 5C; \(^b\) Precursor 23S rRNA molecule shown in Fig. 5B. \(^c\) 23S-5S ITS in *rrnC*. *Value in parenthesis indicates a new corrected size resulting from the examination in this paper. †Percentage identity values in parenthesis indicate percentage identity resulting from inclusion of extra sequences which were not included in the gene annotation for C58.

---

**Fig. 3.** IVS region in 23S rRNA gene predicted by intra-species sequence alignment. Asterisks (*) indicate matching sequences while dashes (-) indicate gaps inserted for alignment. In C58, bold letters indicate part of the 23S rRNA according to the annotation shown in the database. Filled and open arrows indicate the 5' ends of primary and major 23S rRNA molecules revealed in this study (see Fig. 5B, C). Boxed regions indicate sequences unrelated between MAFF301001 and C58.

---

**Fig. 5B**

---

**Fig. 5C**
rrn genes and intervening sequences in Agrobacteria

Fig. 4. Fragmentation of 23S rRNA molecules and removal of IVS. (A, B) Ribosomal RNA of A. tumefaciens MAFF301001 was separated by denaturing gel electrophoresis. Northern blots were subjected to hybridization with ^32P-labeled DNA fragments. Filled arrowheads indicate E. coli 23S and 16S rRNAs. Open arrowheads indicate Agrobacterial rRNA molecules detected by EtBr staining. (C) An illustration of the 23S rRNA fragmentation process. Asterisks (*) represent 5' ends identified in this study (Fig. 5B-D). Filled arrowheads represent cleavage and the gray box indicates the location of IVS.

Fig. 5. Primer extension analysis of 5' ends of 16S and 23S rRNAs and fragments of 23S rRNA. In vitro reverse transcription reactions were carried out using ribosomal RNA as a template with oligonucleotide primers located approximately 80 bp downstream from the predicted 5' ends. (A) 5' end analysis of 16S rRNA using 16SMaf89R; (B) 5' end analysis of 2.9-kb 23S rRNA using 23S5'MafR; (C) 5' end analysis of 2.6-kb 23S rRNA with oligonucleotides 23S283R; and (D) 5' analysis end of 1.3-kb distal fragment of 23S rRNA using 23S4940R. Lefthand lanes (G, A, T, C) are sequence markers generated with the same primer using rrnA DNA. The DNA sequence around the 5' end of each rRNA molecule is presented on the left margin with the position of the end point in boldface marked with an asterisk (*).
Fig. 4B (center panel), the 129-base IVS was absent in the rRNA preparation. However, a 150-base rRNA was detected by Northern hybridization along with the 2.6-kb and 1.3-kb molecules, as shown in Fig. 4B (right panel). These data indicate that the IVS is removed and subsequently 150-base and 2.6-kb fragments are released, and then the IVS disappears probably due to degradation, and 2.6-kb fragments are further cleaved to form two 1.3-kb fragments, as shown in Fig. 4C.

**5' end mapping of 16S and 23S rRNAs, 150-b and and 1.3-kb fragments.** To further characterize the fragmentation events in 23S rRNA molecules and to determine the exact size and primary structure of rRNA of *A. tumefaciens*, 5' ends of the 16S and 23S rRNAs and 1.3-kb rRNA molecules were examined by primer extension analyses, as shown in Fig. 5. Primers complementary to the rRNA fragments were designed using sequences approximately 80 bases downstream from the predicted 5' endpoints as described above. As shown in Fig. 5A, the 5' end of 16S rRNA was positioned 3 bases upstream of the putative 5' end. The resulting size of 16S rRNA was 1495 bases.

As shown in Fig. 5C, the 5' end of the proximal 1.3-kb fragment of 23S rRNA was located 36 bases downstream from the predicted 5' terminus of the putative IVS (Fig. 3). This end was also interpreted as the 5' end of the 2.6-kb 23S rRNA fragment. As shown in Fig. 5D, the 5' end of the distal 1.3-kb 23S rRNA fragment was located 1298 bases downstream of the 5' end of the proximal 1.3-kb molecule. Furthermore, the 5' end of the 150-base fragment was successfully mapped, as shown in Fig. 5B. This position should also be the 5' end of the putative primary 23S rRNA (2.9 kb).

The 5' end data revealed that 16S rRNA is 1495 bases long, while the primary and 2.6-kb 23S rRNAs and the proximal and distal 1.3-kb fragments of 23S rRNA are 2886, 2571, 1298 and 1274 bases long, respectively. We also detected similar-sized rRNA molecules in strain C58C1 by electrophoresis and 5' endpoint determination by primer extension experiments (data not shown).

**DISCUSSION**

In spite of many studies carried out in *Agrobacterium* species using more than 100 rDNA sequences as a tool for classification and phylogenetic purposes and the whole genome sequencing of strain C58 (Wood et al., 2001; Goodner et al., 2001), the *rrn* operons have not been experimentally annotated yet in any *Agrobacterium* species. Early work by Schuch and Loening (1975) showed a precursor molecule of 23S rRNA and its fragments in *Agrobacterium* species. That pioneering work was not pursued and unfortunately was not extended to enable finding of the IVS phenomena. Recently, the presence of IVS has been predicted in *Agrobacterium* strains (Selenska-Pobell and Evgueniieva-Hackenberg, 1995; Zahn et al., 2001; Otten et al., 1996; Selenska-Pobell and Doring, 1998). We conducted detailed characterization of *rrn* by performing intra-species comparisons and biochemical analysis of rRNA molecules in this study.

This study identified 16S rRNA, a precursor of 23S rRNA and fragments of 23S rRNA, and determined their 5' termini. The 129-base sequence of MAFF301001 (not related to the 41-base sequence of C58), which was located 150 bases downstream from the 5' terminus of the 23S rRNA precursor, was missing in rRNA molecules. The presence of 23S rRNA fragments, 150-base, two 1.3-kb fragments and the absence of the 129-base sequence in the rRNAs proved that the IVS phenomenon occurs in 23S rRNA. This is the first information about annotation of IVS as well as rRNA genes in *rrn* operons based on experimental evidence in *Agrobacterium* species. The precise comparative data here would be useful for annotation in other species in the *Rhizobiaceae* family.

The high degree of variation between "operon 4" and the other three operons in C58 suggested that "operon 4" may have a different evolutionary lineage or may have suffered mutation during the course of evolution compared to the remaining three *rrn* operons. The 16S-23S ITS containing *trnVal* and *trnAla* was identical in all the four *rrn* operons in each strain, whereas the ITS differed between MAFF301001 and C58. The 23S-5S ITS of *rrnA* and *rrnC* showed variability in MAFF301001, while only "operon 4" showed such variation in C58. *Agrobacterium* strains have been classified into three biovar groups. Biovar 1 strains have an unconventional genomic organization harboring linear and circular chromosomes in addition to tumor-inducing (Ti) plasmid (for a review, see Suzuki et al., 2004). Both of the two chromosomes harbor rDNAs and chromosomal virulence (*chv*) genes. In contrast, biovar 2 strains contain two circular megareplicons but only one possesses rDNA and *chv* genes (Urbanczyk et al., 2003). The coexistence of two topologically different chromosomes in biovar 1 strains and large chromosomal variations among biovars have attracted our attention regarding their evolution (Jumas-Bilak et al., 1998) and suggest the possibility that the two chromosomes may have originated from two different organisms or from circularization of linear chromosomes or linearization of circular chromosomes during evolution. We expected to reach possible conclusions about chromosome-dependent variations of rDNAs, because the two types of chromosomes harbor *rrn* loci. However, the variation among the four *rrn* operons did not depend on whether the operon was located in the linear or circular chromosome.
Atomic Energy Research Institute to K. Suzuki.

REFERENCES


