Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 6864–6868, July 1993
Microbiology

Intervening sequence with conserved open reading frame in eubacterial 23S rRNA genes

(polymerase chain reaction/spirochete/*Leptospira*)

DAVID RALPH AND MICHAEL MCCLELLAND

California Institute of Biological Research, 11099 North Torrey Pines Road, La Jolla, CA 92037

Communicated by Norman H. Giles, April 1, 1993

ABSTRACT **An intervening sequence (IVS) occurred in the 23S rRNA genes (*rrl*) of some, but not all, strains of four species of the spirochete genus *Leptospira* and was absent from strains in three other species. The IVS varied in size from 485 to 759 base pairs and replaced bases 1224–1245 in both copies of *rrl*. The two ends of each IVS shared 22–35 bases of complementarity that could form a stable double helix. The presence of an IVS correlated with a cleaved mature 23S rRNA that probably results from removal of the IVS without religation. The 3′ site of cleavage was mapped within the inverted repeat of the IVS. An open reading frame of 121–133 amino acids was conserved in the IVS in all four species, oriented so that the sense strand was in the rRNA transcript. When the open reading frames were compared between species, they predicted polypeptides that showed between 51% and 78% amino acid conservation and similar DNA sequence conservation, indicating selection for protein function.**

Intervening sequences (IVSs) that interrupt genes have been found in species in every kingdom, including the nuclear and organellar genomes of eukaryotes and the genomes of archaea and eubacteria and their phage (1–3). Usually these IVSs are spliced from the transcribed RNA using several mechanisms (4, 5). IVSs that are cleaved from the transcribed RNA but not spliced have also been described, particularly in rRNA genes (6). An extreme example of this phenomenon occurs in the trypanosome *Crithidia fasciculata* in which the nuclear large subunit rRNA exists as six fragments encoded by a gene containing five intervening spacers that are processed out in the transcribed RNA but not religated (7).

Eubacterial 23S rRNA genes (*rrl*) have been observed that carry small insertions of about 90–112 bases that are removed by cleavage but not religated (8–10). These IVSs have been found in only some strains within species in a few genera, implying that the IVS may be maintained by lateral transfer between strains but is not necessarily required for host survival.

During a study of restriction site polymorphisms in the *rrl* of various species of the spirochete genus *Leptospira* we observed insertions in *rrl* in the genomes of some strains in four of seven species (11). The 23S rRNAs that carry an IVS are cleaved but not religated, consistent with previous observations in bacteria (8–10). However, we show here that IVSs found in *Leptospira rrl* are much larger than those previously observed, 485–759 bases long, and differ from previously observed IVSs in that they contain a conserved open reading frame (ORF).

MATERIALS AND METHODS

Genomic DNAs. Genomic DNA was prepared from the strains listed in Table 1 using the method of Perolat *et al.* (12).

Oligodeoxyribonucleotide Primers. PCR was performed using primers manufactured by Genosys (The Woodlands, TX):

23S, 1260	CACGGTACCCCTTCGACGGAAGAAAGAACGCTC
23S, 1150	GCAGGAATTCTTTAAAGAGTGCGTAATAGCTCAC
23S, 1432	GGTGTCGACTATGAACCTGCTTCCCATCGACTAC
23S, 240	AACCAGAATTCCGTCAGTAGCGGTGAGCGAA

The underlined sequences are restriction sites used for cloning. These primers were located at base 1253–1284, base 1147–1171, base 220–239, and base 1466–1432 in the *rrl* gene, respectively (GenBank accession no. X14249) (16). In addition, an oligonucleotide, GAACAGTTGGATTTGGCGAACCAA, named Internal IVSa, was designed to hybridize to a sequence within the conserved ORFs of the IVSs and was used as a primer for sequencing the 3′ ends of the sense strand of the IVSs.

PCR, Cloning, and Sequencing. The PCR was performed using recommended conditions (Perkin–Elmer/Cetus) along with 2.5 μCi of [α-^{32}P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq) in a volume of 50 μl. The temperature profile was 1 min at 94°C, 1 min at 50°C, and 2 min at 72°C for 30 cycles. Ten microliters of DNA products was separated by electrophoresis on a 1% agarose gel (FMC). The PCR products derived from strains containing the IVS were cloned into either M13mp18, M13mp19, or pMW-226, which is a modified form of Bluescript (Stratagene), using standard methods (17). For blunt-end ligation into the pMW-226 vector, the ligation mixture contained *Eco*RV to relinearize ligated vector. Transformations were into Epicurian *Escherichia coli* XL1-Blue cells (Stratagene). Single-stranded DNAs for sequencing were rescued using the helper phage VCSM13 (Stratagene). Sequencing reactions were performed with the Sequenase reagent kit (United States Biochemical) using dATP-[α-^{35}S] (Amersham). Electrophoresis and autoradiography used standard methods (17). In all cases, at least two independent clones from each ligation were sequenced.

Southern Blots. Two micrograms of genomic DNAs from a variety of *Leptospira* species was cleaved with *Ava* II using the recommended buffer and conditions (Stratagene). The DNA fragments were resolved on a 1% agarose/TBE (90 mM Tris/64.6 mM boric acid/2.5 mM EDTA) gel. The DNA was transferred to Duralose-UV (Stratagene). A PCR product spanning an IVS was labeled with [α-^{32}P]dCTP (3000 Ci/mmol) using a random priming kit (Stratagene). Hybridization was performed overnight at 55°C using recommended buffers (Boehringer Mannheim). The final wash was in 0.2× SSC/0.1% SDS (1× SSC = 0.15 M NaCl/15 mM sodium citrate) at 55°C. Hybridized DNA products were detected using autoradiography with XAR-5 film (Kodak).

Mapping of the 3′ Cleavage Site in the 23S rRNA. Preparation of rRNA was as described (15). The primer 23S-1432 was end-labeled using polynucleotide kinase [γ-^{32}P]ATP

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IVS, intervening sequence; ORF, open reading frame.

Table 1. Strains of *Leptospira* used in this study

Species/strain*	MRSP profile†	Insert size, bp	Cleaved‡
L. santarosai			
1342 K§	A	485¶	
LT 81	A	485¶	
LT 79	A	≈485	
CI 40	A	≈485	
CZ 188	A	≈485	
TRVL 3214	B	≈485	
TRVL 109873	B	≈485	
L. noguchii			
CZ 214§	A	—	N
Fortbragg	B	518¶	Y
LSU 1945	B	≈518	
L. weillii			
Celledoni§	A	≈550	
Sarmin	B	759¶	
L. borgpetersenii			
Veldrat Batavia 46§	A	—	
Mus 127	A	—	
Sari	A	—	
M 84	A	—	
Arborea	A	—	
Jules	A	—	
Njenga	A	—	
Perepelicyn	B	—	
1627 Burgas	C	≈683	Y
Sorex Jalna	C	≈683	
Hardjo bovis	D	683¶	Y
L. interrogans			
RGA§	A	—	N
M 20	A	—	
Djasiam	A	—	
Hebdomadis	A	—	
Jalna	A	—	
Jez-bratislava	A	—	
Gurung	A	—	
Kremastos	A	—	
Paidjan	A	—	
Bangkinang I	B	—	
Hardjoprajitno	B	—	N
Jones	B	—	
Lai	B	—	
Moores	B	—	
L. meyeri			
ICF	B	—	
267-1348	B	—	
L. kirchneri			
3522 C§	A	—	
Moskva	A	—	N
Kipod 179	A	—	
Kamituga	B	—	
Ndambari	B	—	
LT 1014	C	—	
Wumalasena	C	—	
L. biflexa			
Patoc I§		—	

*Strain and species assignments are from refs. 12, 13, and 14. For serovar assignments, see table 1 in ref. 11.
†This column lists the 16 different rRNA gene profiles of mapped restriction site polymorphisms (MRSP) that constitute an intraspecific classification (11).
‡As determined here or by Hsu *et al.* (15).
§Type strain.
¶The IVS was sequenced from these strains. A dash (—) indicates an insert was not detected in the PCR product spanning the insertion region.

(3000 Ci/mmol) and then used to prime reverse transcription with Moloney reverse transcriptase (Stratagene) in standard buffer and conditions (17). Extension products were resolved by electrophoresis on 6% polyacrylamide/7 M urea gels. The gels were fixed in 10% glacial acetic acid/12% methanol, dried, and autoradiographed using XAR-5 film (Kodak).

RESULTS AND DISCUSSION

An IVS in the 23S rRNA Genes of Some *Leptospira*. The 23S rRNA genes (*rrl*) from 47 serovar reference strains in seven species of the spirochete genus *Leptospira* were PCR amplified and strains were grouped into species using mapped restriction site polymorphisms (11). Unusually long *rrl*s were observed in all strains examined in two species and in some strains of two other species. A total of 14 of 47 isolates tested contained an insertion in *rrl* (Table 1). The insertion was localized to a region between 1200 and 1280 bp relative to the *L. interrogans* 23S rRNA sequence (GenBank accession no. X14249) using restriction digests of PCR products. One strain each from the four species that carried the insertion was examined further. The region of *rrl* spanning the insertion was cloned in both orientations in M13 vectors or pMW-226 (a Bluescript-derived vector; Stratagene) and sequenced. Complete sequences spanning the insertions were obtained. Insertions ranging from 485 to 759 bp were found in the unusually long *rrl*s, replacing the sequence between bases 1224 and 1245 in other *Leptospira* (Fig. 1). This location brackets the site of insertion of a sequence of about 90 bp observed in the genomes of some strains of the enterobacterial genera *Salmonella* (8) and *Yersinia* (9) and an insertion of 112 bp in a strain of *Actinobacillus* (10), all of which are phylogenetically distant from *Leptospira*. The IVSs in the *rrl*s from strains in four *Leptospira* species did not have sequences or implied secondary structures characteristic of group I (4) or group II introns (18).

The Ends of the IVS May Combine to Form a Stem–Loop. An RNA folding program (19, 20) detected large inverted repeats that flank the core of each IVS very near their boundaries with the surrounding 23S rRNA sequences. The shorter IVSs in 23S rRNAs from *Salmonella* (8), *Yersinia* (9), and *Actinobacillus* (10) also have extensive inverted repeats. The terminal inverted repeats in the *Leptospira* IVSs varied between 23 and 31 bp in length with only one to three mismatches. Such sequences could potentially form helices with considerable stability (Fig. 1). The mismatches within these putative helices may be RNase III processing sites and, indeed, the mapped 3′ cleavage site is near a mismatch in the proposed stem structure (see below). RNase III was shown to cut the terminal inverted repeat at both ends to excise the IVS from the mature 23S rRNA of *Salmonella in vitro* (8). All four IVSs from *Leptospira* also had further internal secondary structures though none was as potentially stable as the ones at the ends of the insertion.

The 23S rRNAs that Carry the IVS Are Cleaved. In *Salmonella, Yersinia*, and *Actinobacillus* the IVSs in the *rrl* are cleaved at both ends of the insert with a few nucleotides from the IVS remaining in the 23S rRNA-derived products (8–10). The remaining few bases of IVS sequence are not religated or spliced, resulting in a mature and functional 23S rRNA composed of two molecules. Similarly, the 23S rRNA was fragmented into 14S and 17S products in some but not all strains of *Leptospira* (15). Cleavage of the 23S rRNA correlated with the presence of an IVS (Table 1). To confirm this observation, a cleavage site was mapped in *L. noguchii* fortbragg and *L. weillii* sarmin using primer extension by reverse transcription (see *Materials and Methods*). The site of cleavage is in the stem structure near the middle of the proposed terminal helix in the IVS and about 15 bases from the 3′ end of the IVS (Fig. 1). Strains of *L. interrogans* and

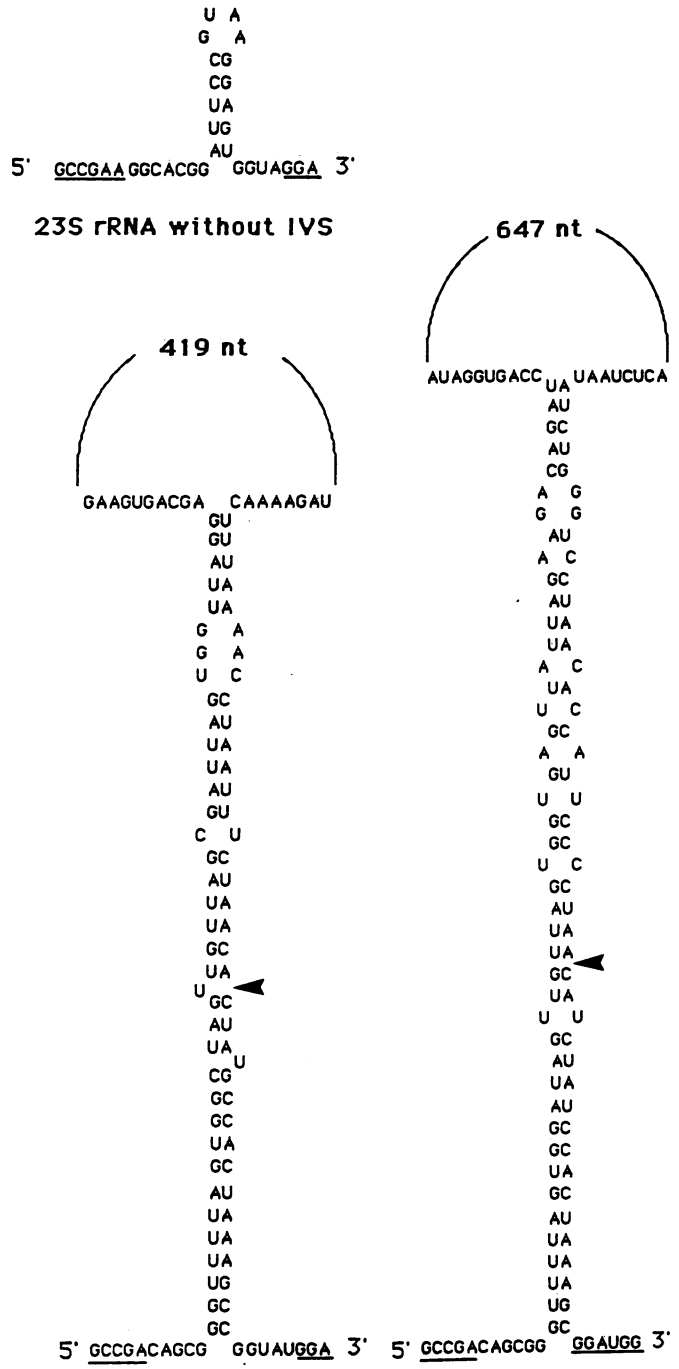

FIG. 1. DNA secondary structure at the junction of *rrl* and the IVS. The sequences presented are for *L. noguchii* fortbragg (serovar and strain uncertain) and *L. weillii* sarmin Sarmin. The complete sequences of the four IVSs determined in this study have been deposited in GenBank (accession nos. L14263–L14266). The site of cleavage at the 3′ end of the IVS is shown by an arrowhead. The probable cleavage at the 5′ end of the IVSs has not been mapped. The 23S RNA without an IVS is from *L. interrogans* canicola Moulton (16) and spans base 1219–1251 in the latter sequence. The 23S rRNA sequence flanking each IVS is underlined.

L. noguchii that did not have an IVS were not cleaved in this region (data not shown). It is probable that cleavage in the *Leptospira* IVS also occurs in the 5′ end, adjacent to the mapped site, as it does in *Salmonella* (8), *Yersinia* (9), and *Actinobacillus* (10). Cleavage would then result in a 15-bp stem that would be quite stable and, incidentally, would explain why the cleaved *Leptospira* 23S rRNA pieces were still bound to each other during electrophoresis unless heated to 65°C prior to loading (15).

Nucleotides 1231–1245 of *rrl* would normally form a stem that is 5 bp long with a four-base loop (16, 21) (Fig. 1). This stem–loop structure is also a feature of 23S rRNAs in other bacterial taxa, although its length varies widely. In those *Leptospira* strains that contained the IVS, this conserved stem was replaced by the inverted repeats of the IVS, which replaced the sequence between nucleotides 1224 and 1245 in the rRNA. Similarly, the IVS observed in some strains of *Salmonella, Yersinia*, and *Actinobacillus* replaced about 20 bases of sequence in part of the same stem–loop of *rrl*. However, the situation in *Leptospira* is unique because not only is this entire stem replaced by IVS sequences but also the sequence between nucleotides 1224 and 1231, which is usually involved in an interaction that is highly conserved among 23S rRNAs, is disrupted. This sequence normally participates in a highly conserved helix by base pairing with a sequence, 5′-CCGT(C)GTTT, about 160 bases earlier in the rRNA. The IVS replacing bases 1224–1231 does not participate in the predicted stem in the IVS formed by the large inverted repeats inside the IVS (Fig. 1) and might, therefore, maintain some of the conserved helix despite only moderate complementarity. The 3′ end of the IVS at nucleotide 1245 is near the end of the stem in the normal 23S rRNA and should not disrupt 23S rRNA structure.

The IVS Occurs Only in *rrl*. A Southern blot was performed at moderate stringency against a variety of *Leptospira* genomic DNAs using the cloned IVS and 140 bp of flanking *rrl* sequence from a *L. noguchii* strain (Fig. 2). The 1186-bp *Ava* II fragment in lanes 1–4, 6, and 12 represents the 23S rRNA lacking an IVS. Lanes 5–11 contain *Ava* II products that add up to 1186 bp plus the size of the IVS. The extra products in lanes 5 and 11 were due to *Ava* II sites in the IVS. The blots indicate that copies of the IVS were not found elsewhere in the genome nor were they found in strains that did not carry the insertion in *rrl* (Fig. 2). Furthermore, *L. interrogans*, and presumably all other *Leptospira*, carry two copies of *rrl* (22),

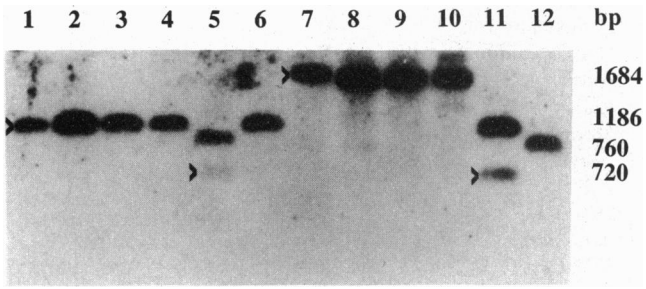

FIG. 2. Southern hybridization using IVS DNA as a probe. Genomic DNAs were cleaved with *Ava* II, which cuts within *rrl* on both sides of the IVS insertion. Fragments were resolved and hybridized as described in the text. The probe used was a PCR fragment spanning the *rrl* of *L. noguchii* serovar fortbragg from base 1147 to base 1284, plus the 518-bp IVS. The nucleotide sequence lengths determined by sequencing of the products are indicated by ">". Lanes 5 and 7–11 contain genomic DNAs from strains that carry an IVS. The other lanes are from strains that do not have the IVS. Lane 1, *L. interrogans*, serovar icterohaemorrhagiae, strain RGA (type strain); lane 2, *L. interrogans* copenhageni, M 20; lane 3, *L. kirshneri* cynopteri 3522C; lane 4, *L. borgpetersenii* ballum Mus 127; lane 5, *L. borgpetersenii* balcanica 1627 Burgas; lane 6, *L. noguchii* panama CZ 214 (type strain); lane 7, *L. noguchii* fortbragg, strain uncertain; lane 8, *L. santarosai* shermani 1342 K (type strain); lane 9, *L. santarosai* bakeri LT 79; lane 10, *L. weillii* celledoni Celledoni (type strain); lane 11, *L. weillii* sarmin Sarmin; lane 12, *L. biflexa* patoc Patoc I^T.

so the blots also indicate that both genes carry an identical insertion in those strains with an insertion. This latter conclusion was also supported by the size of PCR products using 23S rDNA (rRNA-encoding DNA)-specific primers and by restriction digestion of the uncloned PCR amplification products, which demonstrated that all *rrl*s within the genomes of the individual strains examined in this study were of a single size class (ref. 11; data not shown). In contrast, in some strains of *Salmonella* and *Yersinia*, the analogous but smaller IVSs in the *rrl*s did not occur in all copies of the gene (8, 9).

An ORF in the IVS Is Conserved Among Species. Another significant feature shared by all four of the sequenced *Leptospira* IVSs was the presence of ORFs that potentially encode proteins of 121–133 amino acids in length (Fig. 3). These ORFs were oriented so that the reading frame was present in the 23S rRNA primary transcript and would therefore be one the most abundantly transcribed RNAs in the cell with the potential to be translated. When these ORFs were compared between IVSs derived from the four different species of *Leptospira*, the amino acid sequences were found to be similar and collinear. At numerous positions the predicted amino acid changes were conservative. Amino acid identity of the ORFs varied from 49% to 78% in pairwise comparisons and this degree of amino acid conservation of the ORFs was consistently similar to the amount of DNA conservation in the same region (Table 2). Because amino acid sequences would be predicted to diverge at least twice as fast as nucleotide sequences in the absence of selection, this conservation strongly suggested that the amino acid sequence was under selective pressure. Furthermore, of 46 conserved amino acid positions with possible codon redundancy, 33 were encoded by multiple codons. For example, leucine at amino acid 42 is encoded by four different codons in the four IVSs. Comparison with the DNA to DNA hybridization studies (13, 14) shows that the degree of sequence similarity between these IVS-derived ORFs is typical of the degree of sequence similarity (50–86%) observed between the genomes of these species of *Leptospira*. Functionally homologous protein coding genes in these species would, therefore, be expected to vary by about the same degree as was observed with the IVS ORFs. Finally, outside the region of the ORF the conservation of IVS sequence between species was so poor that an alignment could not be achieved with confidence. From all these data, we predict that the ORFs are expressed and that they encode polypeptides that have functions that are under evolutionary constraint.

The DNA and amino acid sequences of the *Leptospira* IVSs were compared with the GenBank v70 and Protein Identification Resource v27 data base using FASTA software but no matches were found. Although the function of the predicted products of the IVS ORFs remains to be determined, several possibilities can be imagined. For instance, intron-encoded proteins required for DNA cleavage during lateral transfer of introns have been observed in chloroplast *rrl*s (e.g., ref. 23). Skurnik and Toivanen (9) have suggested that the clean removal of most of the IVS in the *Yersinia* 23S rRNA may result in a structure that is resistant to bacteriocins (24) directed against the rRNA. If so, the ORF might encode the bacteriocin that acts on the rRNA at that site. However, this possibility is made less likely by the fact that the IVSs in the 23S rRNAs of other genera do not encode a protein (8–10) and removal can be achieved by cleavage with RNase III (8).

Another possible function of the IVS ORF is that it is an essential part of the IVS as a mobile genetic element. There is considerable precedent for bacterial mobile genetic elements conveying a selective advantage to their hosts (25) and for proteins encoded by mobile genetic elements being required for their observed mobility (25, 26), and inverted repeats are known to flank a number of mobile genetic elements (25).

```
                                                                  20
sant  M A G F R N I D D E D V E R R A Y G L S
nogu  M V S F K K I E D E D V E K R A Y D L S
weil  M A G L K R I E D E D V E N R A Y A L S
borg  M D K I R K Y E D L E V Y R K S Y L V S

                                                                  40
sant  L K I H K L S L E F P K E E Q S D L G N
nogu  L K I H K F S L S L P K E E Q S D L A N
weil  L K I D K I S L E M P K E E Q L D L A N
borg  L E I H K K T L E F P K E E Q Y G L A S

                                                                  60
sant  Q L R R S S K S V C S N E A E G E A K Q
nogu  Q L R R S S K S V C S N E A E G E A K Q
weil  Q L R R S S K S I C S N E A E G E A K Q
borg  Q L R N S S K S I C G N I A E G E A K Q

                                                                  80
sant  S F S A K E E R K Y L I I G L G S S D E
nogu  N F S V K E E R K Y L I I A L G S S D E
weil  N F S K R E E K K Y L I I A I G S A N E
borg  S Q S K Q E E R R F L S M T I G S A D E

                                                                 100
sant  S I L W L R Y C K D L G Y L S E K V W L
nogu  S M L W L R Y C K D L G Y L S E S E S D
weil  S A L W L R Y S K D L G Y I T E D Q Y S
borg  T K V W L N Y A K D L G Y I T E S L Y T

                                                                 120
sant  E M T N E Y L E I S K X L A V L A K K S
nogu  E W T E E Y L Q V S K M L A R L A K L Q
weil  E W S D E C L Q I A K I L S K L A Q I Q
borg  S W K G T Y V D I A K M L S G L Y K S W

sant  S N *
nogu  D S K F L R Q T V N R Q Y *
weil  A V S E *
borg  E *
```

FIG. 3. Conceptual amino acid sequences of the ORF in the intervening sequence of some *Leptospira rrl*s. Strains examined were "sant," *L. santarosai* atlantae LT 81; "nogu," *L. noguchii* fortbragg (serovar and strain uncertain); "weil," *L. weillii* sarmin Sarmin; and "borg," *L. borgpetersenii* hardjo, Hardjo bovis/Sponselee. Amino acids conserved between three or more species are underlined. An asterisk (*) indicates termination codon.

Distribution of IVSs in *Leptospira* and Other Eubacteria. The distribution of IVSs in only some strains of some species *Leptospira* is characteristic of a mobile genetic element. However, the sample of IVSs that we have sequenced does not provide direct evidence of lateral transfer between species or within species.

An alternative is that an IVS was present in the last common ancestor of all four species of *Leptospira* and that it has been lost from some of the lineages. The absence of these IVSs in many lineages would then indicate that there was

Table 2. DNA and amino acid sequence conservation in the ORF

Species	% conservation*		
	L. noguchii	*L. weillii*	*L. borgpetersenii*
L. santarosai			
DNA	77	68	56
Protein	78	66	50
L. noguchii			
DNA		68	56
Protein		69	49
L. weillii			
DNA			59
Protein			51

See Fig. 3 for the strains compared.
*% sequence conservation in the first 363 bases or first 121 amino acids.

little or no selective advantage in having one of these IVSs in most circumstances. Nevertheless, there was selection for the ORFs in those organisms that maintain an IVS, as indicated by the conservation of the ORFs. The requirements of the strain in different habitats or host ranges might then explain the distribution of the IVS.

Sequencing of more IVSs and of other genes will be required to perform a more extensive phylogenetic comparison. Lateral transfer of these IVSs, if demonstrated, would be the first evidence of gene transfer in spirochetes. Such a gene transfer mechanism could prove technologically useful.

Previous publications have shown the removal of IVSs of about 100 bases from 23S rRNA without religation in *Salmonella* (8), *Yersinia* (9), and *Actinobacillus* (10). Fragmentation of the 23S rRNA has been observed in the diverse eubacteria *Rhodobacter, Anacystis*, and *Micrococcus* (referred to in ref. 8), and *Brucella* and *Agrobacterium* (27), although it is yet to be established whether these cases involve an IVS. A potential IVS was observed in the chloroplast *rrl* of *Chlorella ellipsoidea* (28). This IVS is 243 bp and was found in the equivalent position as those described here, although whether fragmentation occurs in the rRNA is yet to be established. Taken together these observations suggest that the phenomenon of an IVS in the large subunit rRNA that is transcribed and cleaved, but not religated, could be widespread among eubacteria and organelles, as it is in eukaryotes (7). Here we have added the observation of such an IVS, which, in addition, carries a conserved ORF. The prevalence of IVSs with conserved ORFs and the function for the encoded protein are yet to be determined.

We particularly thank Guy Baranton and Phillipe Perolat for supplying strains and for their enthusiastic collaboration, without which this work would not have been possible. We thank John Welsh, Bruno Sobral, K. K. Wong, and Cibra Sinquing for critical reading of the manuscript. D.R. was supported by California Institute of Biological Research set-up funds. This work was supported in part by National Institutes of Health Grants AI34829 and HG00456 to M.M. and by Grant AI32644 to John Welsh.

1. Belfort, M. (1990) *Annu. Rev. Genet.* **24**, 363–385.
2. Reinhold-Hurek, B. & Shub, D. A. (1992) *Nature (London)* **357**, 173–176.
3. Perler, F. B., Comb, D. G., Jack, W. E., Moran, L. S., Qiang, B., Kucera, R. B., Benner, J., Slatko, B. E., Nwankwo, D. O., Hempstead, S. K., Carlow, C. K. S. & Jannasch, H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5577–5581.
4. Cech, T. (1990) *Annu. Rev. Biochem.* **59**, 543–568.
5. Jacquier, A. (1990) *Trends Biochem. Sci.* **15**, 351–354.
6. Raué, H. A., Klootwijk, J. & Musters, W. (1988) *Prog. Biophys. Mol. Biol.* **51**, 77–129.
7. Spencer, D. F., Collings, J. C., Schnare, M. N. & Gray, M. W. (1987) *EMBO J.* **6**, 1063–1071.
8. Burgin, A. B., Parodos, K., Lane, D. J. & Pace, N. R. (1990) *Cell* **60**, 405–414.
9. Skurnik, M. & Toivanen, P. (1991) *Mol. Microbiol.* **5**, 585–593.
10. Haraszthy, V. I., Sunday, G. J., Bobek, L. A., Motley, T. S., Preus, H. & Zambon, J. J. (1992) *J. Dent. Res.* **71**, 1561–1568.
11. Ralph, D., McClelland, M., Welsh, J., Baranton, G. & Perolat, P. (1993) *J. Bacteriol.* **175**, 973–981.
12. Perolat, P., Grimont, F., Regnault, B., Grimont, P. A. D., Fournie, E., Thevenet, H. & Baranton, G. (1990) *Res. Microbiol.* **141**, 159–171.
13. Yasuda, P. H., Steigerwalt, A. G., Sulzer, K. R., Kaufmann, A. F., Rogers, F. & Brenner, D. J. (1987) *Int. J. Syst. Bacteriol.* **37**, 407–415.
14. Ramadass, P., Jarvis, B. D. W., Corner, R. J., Penny, D. & Marshall, R. B. (1992) *Int. J. Syst. Bacteriol.* **42**, 215–219.
15. Hsu, D., Pan, M.-J., Zee, Y. C. & LeFebre, R. B. (1990) *J. Bacteriol.* **172**, 3478–3480.
16. Fukunaga, M., Horie, I. & Mifuchi, I. (1989) *Nucleic Acids Res.* **17**, 2123.
17. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
18. van der Veen, R., Arnberg, A. C., van der Horst, G., Bonen, L., Tabak, H. F. & Grivell, L. A. (1986) *Cell* **44**, 225–234.
19. Zuker, M. & Steigler, P. (1981) *Nucleic Acids Res.* **9**, 133–148.
20. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
21. Gutell, R. R. & Fox, G. E. (1988) *Nucleic Acids Res.* **16**, r175–r269.
22. Fukunaga, M. & Mifuchi, I. (1992) *J. Bacteriol.* **171**, 5763–5767.
23. Marshall, P. & Lemieux, C. (1991) *Gene* **104**, 241–246.
24. Luria, S. E. & Sluit, J. L. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. Neidhart, F. C. (Am. Soc. Microbiol., Washington, DC), pp. 104–125.
25. Berg, D. & Howe, E. (1989) *Mobile DNA* (Am. Soc. Microbiol., Washington, DC).
26. Dujon, B., Belfort, M., Butow, R. A., Jacq, C., Lemieux, C., Perlman, P. S. & Vogt, V. M. (1989) *Gene* **82**, 115–118.
27. Hsu, D., Zee, Y. C., Ingraham, J. & Shih, L. M. (1992) *J. Gen. Microbiol.* **138**, 199–203.
28. Yamada, T. & Shimaji, M. (1987) *Curr. Genet.* **11**, 347–352.