The use of 16S rDNA sequence analysis to investigate the phylogeny of Leptospiraceae and related spirochaetes

JOHN V. HOOKY,1* JULIA BRYDEN2 and LAURENCE GATEHOUSE2

1PHLS Leptospira Reference Laboratory, FAO/WHO Collaborating Centre, County Hospital, Hereford HR1 2ER, UK
2Department of Biological Sciences, The Science Laboratories, University of Durham, Durham DH1 3LE, UK

(Received 18 May 1993; revised 5 July 1993; accepted 15 July 1993)

The 16S rDNA sequences from 15 Leptospiraceae were determined by automated PCR-directed cycle sequencing. Nucleotide comparisons, including those from published sequences for Leptospira canicola Moulton and Serpulina spp., were used to construct phylogenetic trees. Serpulina hyodysenteriae and S. innocens were related to each other but were distinct from the Leptospiraceae comprising Leptospira parva incertae sedis (Turneria parva H), Leptonema illini and Leptospira spp. The pathogenic and the saprophytic leptospires were distinct and separated from each other. Leptospira inadai occupied an intermediate position between the two forms. The pathogens formed three groups. Group I was represented by L. interrogans sensu stricto and L. kirschneri, Group II by L. weilii, L. borgpetersenii and L. santarosai, and Group III comprised L. noguchii and L. meyeri. The saprophytic species, L. wolbachii and L. biflexa sensu stricto shared about 99% sequence similarity. The freshwater isolates were distinct from the marine isolate L. biflexa sensu lato ancona Ancona Porto.

Introduction

The Leptospiraceae, comprising Leptospira and Leptonema, are motile, flexible helical aerobic spirochaetes. Some are pathogenic for man and animals and others are free-living in soil, fresh- and salt-waters (Johnson & Faine, 1984). Their classification remains problematic and two different classifications co-exist, one based on serological affinities and the other on genetic relatedness (Kmety & Dikken, 1988; Yasuda et al., 1987). Traditionally, two species Leptospira interrogans sensu stricto (the pathogens) and L. biflexa sensu lato (the so-called 'saprophytes') could be delineated on the basis of their DNA base composition and a limited number of phenotypic characteristics (Johnson & Faine, 1984), and each in turn was arranged into the basic taxon of serovar; L. interrogans has 202 serovars and L. biflexa 65 (Kmety & Dikken, 1988). Recently eight new species, L. borgpetersenii, L. inadai, L. meyeri, L. noguchii, L. santarosai, L. weilii, L. wolbachii (Yasuda et al., 1987) and L. kirschneri (Ramadass et al., 1992) were described on the basis of DNA:DNA homology studies. Similarly, Ramadass et al. (1990) revealed that the L. biflexa sensu lato 'complex' comprised at least six DNA homology groups.

Comparative sequencing of small subunit rRNA is now the method of choice for constructing higher level classifications (Woese, 1987). Oligonucleotide cataloguing and, more recently, reverse transcriptase sequencing of 16S rRNA showed that the spirochaetes were a phylogenetically distinct group subdivided into the Spirochaetaceae (Borrelia, Cristispira, Serpulina, Spirochaeta and Treponema) and the Leptospiraceae [Leptospira and Leptonema (Paster et al., 1984, 1991)].

In view of the limited amount of information available with regard to the evolutionary relatedness of leptospires, the intent of this study was to establish more fully the phylogenetic relationships between members of the family Leptospiraceae on the basis of 16S rDNA sequence analysis.

Methods

Bacterial strains and culture. Fifteen strains of Leptospiraceae (Table 1) were cultured in EMJH broth (Ellinghausen & McCullough, 1965)
containing no DNA served as a negative control.

was added to the DNA. Thermal cycling
chromosomal DNA, the reaction mix comprising 10
0-01
0.1% gelatin; EIA grade, Bio-Rad), 200 pm each of dATP, dCTP,
dGTP and dUTP (Pharmacia), 100 pm of each primer. Glycerol at 10
(w/v), 2.5 U AmpliTaq DNA polymerase (ILS Ltd) and sterile
oligomers purified according to the manufacturer's recommendations.

aerobic bacteria by overnight culture on 8
using a DNA-synthesizer (model 381, Applied BioSystems) and the
30 and 37 °C.

16s rDNA. DNA was amplified by
PCR using a primer pair specific for 16s rDNA
reaction buffer (500 mM-KC1, I00 mM-Trk/HCl, pH 8.3, 35 mM-MgCl,,
(dye (ATCC 43289))

Leptonema illini**

L. biflexa sensu lato*

Bacterial strains examined and their rDNA sequence accession numbers

<table>
<thead>
<tr>
<th>Species</th>
<th>Serovar Reference Strain*</th>
<th>Source†</th>
<th>Accession No.‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. borgpetersenii§</td>
<td>javanica Veldrat Batavia 46 (ATCC 43292)¶</td>
<td>W. J. Terpstra (Royal Tropical Institute, Melibergdreef, Amsterdam, The Netherlands)</td>
<td>Z12817</td>
</tr>
<tr>
<td>L. interrogans sensu stricto§</td>
<td>icterohaemorrhagiae RGA¶</td>
<td>W. J. Terpstra</td>
<td>Z12649</td>
</tr>
<tr>
<td>L. noguchii§</td>
<td>panama CZ 214 (ATCC 43288)¶</td>
<td>W. J. Terpstra</td>
<td>Z12635</td>
</tr>
<tr>
<td>L. santarosai§</td>
<td>shermani 1342K (ATCC 43286)¶</td>
<td>W. J. Terpstra</td>
<td>Z12637</td>
</tr>
<tr>
<td>L. weilli§</td>
<td>celledoni Celledoni (ATCC 43285)¶</td>
<td>A. F. Kaufmann (Centres for Disease Control, Atlanta, GA, USA)</td>
<td>Z12638</td>
</tr>
<tr>
<td>L. kirschneri§</td>
<td>cynopteri 3522 C¶</td>
<td>B. Cacciapuoti (Instituto Superoire di Sanita, Roma, Italy)</td>
<td>Z12821</td>
</tr>
<tr>
<td>L. inadai§</td>
<td>lyme (ATCC 43289)¶</td>
<td>B. Cacciapuoti</td>
<td>Z12648</td>
</tr>
<tr>
<td>L. biflexa sensu stricto**</td>
<td>patoc Patoc 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. meyeri**</td>
<td>ranarum ICF (ATCC 43287)¶</td>
<td>W. J. Terpstra</td>
<td>Z12639</td>
</tr>
<tr>
<td>L. wolbachii**</td>
<td>edc Biflexa CDC (ATCC 43284)¶</td>
<td>B. Cacciapuoti</td>
<td>Z12636</td>
</tr>
<tr>
<td>L. para incertae sedis**</td>
<td></td>
<td>K. Hovind-Hougen (National Veterinary Laboratory, Copenhagen, Denmark)</td>
<td>Z12632</td>
</tr>
<tr>
<td>Leptonema illini**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. biflexa sensu lato*</td>
<td>anconca Anconca Porto canela Canela inequitaia inequitaia</td>
<td>B. Cacciapuoti</td>
<td>Z12629</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Z12631</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Z12633</td>
</tr>
</tbody>
</table>

† Designation based on serological studies (Kmety & Dikken, 1988).
‡ Accession numbers were obtained from the EMBL data library. The accession number for canicola Moulton was X17547 (Fukunaga et al., 1990) and those for S. hyodysenteriae and S. innocens were M57741 and M57744, respectively (Stanton et al., 1991).
§ Reference strains assigned to a species on the basis of DNA hybridization (hydroxyapatite method; Yasuda et al., 1987).
¶ Reference strains assigned to a species on the basis of DNA hybridization (slot-blot method; Ramadass & Marshall, 1990; Ramadass et al., 1992).
** Reference strains assigned to a species on the basis of DNA hybridization (hydroxyapatite method; Yasuda et al., 1987).
* Species considered to be pathogenic for man and animals (Yasuda et al., 1987).
† Reference strains assigned to a species on the basis of DNA hybridization (slot-blot method; Ramadass & Marshall, 1990; Ramadass et al., 1992).
** Species considered to be saprophytic (Yasuda et al., 1987; Johnson & Faine, 1984).

at 30 °C for 7–10 d and checked for the presence of contaminating aerobic bacteria by overnight culture on 8% (w/v) horse-blood agar at 30 and 37 °C.

Isolation of chromosomal DNA. Chromosomal DNA was prepared by the method of Hookey & Palmer (1991).

Preparation of PCR-amplified 16S rDNA. DNA was amplified by PCR using a primer pair specific for 16S rDNA [pU (34–53): Hookey, 1992; pH˚: Edwards et al., 1989]. Primer oligonucleotides were made using a DNA-synthesizer (model 381, Applied BioSystems) and the oligomers purified according to the manufacturer's recommendations. Reactions were performed on a Perkin-Elmer 480 Thermal Cycler in sterile thin-walled tubes that contained approximately 100 ng (1–2 µ) chromosomal DNA, the reaction mix comprising 10 µ of 10 × PCR reaction buffer (500 mm-KCl, 100 mm-Tris/HCl, pH 8.3, 35 mm-MgCl, , 0.01% gelatin; EIA grade, Bio-Rad), 200 µ each of dATP, dCTP, dGTP and dUTP (Pharmacia), 100 pm of each primer. Glycerol at 10% (w/v), 2.5 U AmpliTaq DNA polymerase (ILS Ltd) and sterile 'Millipore' water were added to a final volume of 100 µ. The reaction mix was overlayed with 150 µ mineral oil (ILS Ltd). A reaction mix containing no DNA served as a negative control.

Following denaturation of the DNA template at 95 °C for 5 min, the temperature was lowered to and held at 80 °C while the reaction mix was added to the DNA. Thermal cycling (×30) then proceeded at 95 °C for 40 s, 55 °C for 25 s and 72 °C for 3 min. A final step at 72 °C for 10 min was included.

The amplified DNA was extracted once with an equal volume of phenol/chloroform, precipitated in the presence of 0.1 vols 3 M-sodium acetate (pH 5.2) and 3 vols absolute ethanol (AnalaR; BDH), and redissolved in 25 µl Millipore water.

Purification of 16S rDNA product. The amplified 16S rDNA was purified prior to DNA sequencing according to standard protocols (Sambrook et al., 1989). Briefly, unwanted primers, deoxyribonucleotides and DNA polymerase were separated from the amplification product by electrophoresis through a 1.0% (w/v) agarose gel (MB grade; Bio-Rad) at 100 mA onto activated DEAE-cellulose membranes (NA-45; Schleicher & Schuell).

Automated PCR-directed cycle sequencing. The purified double-stranded DNA template was sequenced using a Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied BioSystems) on an ABI model 373A automated fluorescent DNA sequencer according to the manufacturer's instructions. The primers used were pC, pD, pD’, pH’ (Edwards et al., 1989), pU (Hookey, 1992), p4 [5'...TAA GGG TTG CGC TCG TTG...3'] (1132–1115) and pE4 [5'...CGC AAC GAG CGC AAC CCT...3'] (1113–1130).

The thermal cycle sequencing programme used on the Pharmacia-LKB thermal cycler was 96 °C for 30 s, 50 °C for 40 s, and a 4 min
extension at 60 °C for 25 cycles. Following amplification, excess DyeDecoy Terminators were removed from the sequencing reaction by spin-column centrifugation according to the manufacturer’s instructions (Bio-Spin 30 chromatographic column; Bio-Rad), ethanol-precipitated, dried and redissolved in 4 µl loading buffer. DNA fragments were separated on conventional 0.4 mm, 6.0 % DNA gel loading buffer.

Analysis of sequence data. The primary sequence data have been deposited with the GenBank/EMBL data libraries and the accession numbers are given in Table 1. Sequences were aligned using the MULTALIGN program (Corpet, 1988; Cherwell Scientific Publishing Ltd, Oxford, UK). The 11 Leptospira spp., Leptonema illini, four reference serovars and two related spirochaetes were checked against the secondary structure model of Neefs et al. (1991). Those regions of hypervariability and base positions that could not be aligned unambiguously were omitted from the final alignment (about 140 nucleotides similar to those omitted by Paster et al., 1991). A total of 1339 bases comprised the final alignment; this is available from the authors on request.

Distance matrices. Distance matrices were constructed from bootstrapped data by DNADIST contained in the Phylogenetic Inference Package, PHYLIP 3.4 (J. Felsenstein, University of Washington, USA) and user trees computed using neighbour-joining (NEIGHBOR; Saitou & Nei, 1987), and Fitch–Margoliash (FITCH; Fitch & Margoliash, 1967) using either the Jukes & Cantor (1969) correction for multiple substitutions or the maximum likelihood method (DNAML) employing the empirical frequencies of the bases as given in the input data. Majority-rule consensus trees were computed from the user trees using the CONSENSE program.

The bootstrap percentages quoted for a species or serovar in the legend to Fig. 1 are the percentage times a taxa at that node occurred. The scale bar represents 0.02 substitutions per sequence position (Kₑₑₑₑ). The bootstrap percentages quoted are the percentage times a taxa at that node occurred. The scale bar represents 0.02 substitutions per sequence position (Kₑₑₑₑ).

Fig. 1. Unrooted consensus Jukes & Cantor (1969) correction and Fitch & Margoliash (1967) distance tree showing the position of Leptospiraceae and related spirochaetes. Bootstrap percentages for the following were: L. interrogans sensu stricto and canicola Mouton, 78 %; L. kirschneri, 98 %; L. noguchii and L. meyeri, 34 %; L. santarosai, 67 %; L. weili and L. borgpeterseni, 75 %; ancona Ancona Porto, 100 %; jequitaia Jequitaia, 39 %; canela Canela, 27 %; L. biflexa sensu stricto and L. wolbachii, 96 %; L. parva incerta sedis, 56 %; Leptonema illini, 100 %; E. coli, 97 %; S. hydysenteriae and S. innocens, 100 %. The bootstrap percentages quoted are the percentage times a taxa at that node occurred. The scale bar represents 0.02 substitutions per sequence position (Kₑₑₑₑ).

Results and Discussion

Sequencing generated between 1451 and 1374 bases (97–92 %) of 16S rDNA sequence for the 15 strains studied. The matrix of evolutionary distance (Kₑₑₑₑ) values are given in Table 2.

Phylogenetic trees based on sequence data must be interpreted with caution (Felsenstein, 1988). In this study, bootstrapped data, out-grouping to E. coli and the statistical comparisons of topologies indicated that the
branching order was reliable and that the consensus tree (Fig. 1) may be a good reconstruction of the relationships between the *Leptospiraceae*. The overall topology between the *Serpulina* spp., *Leptonema illini*, *Leptospira interrogans* and *L. biflexa* was identical to that given by Paster et al. (1991).

The consensus trees summarize the evolutionary relationships between taxa examined and their relative positions (Fig. 1). The aetiological agent of swine dysentery, *Serpulina hyodysenteriae* and the saprophyte *S. innocens* were distinct species (Stanton et al., 1991; Fig. 1; Table 2). Their exact phylogenetic position relative to the *Leptospiraceae* was not resolved by the current data and their position as an in-group in Fig. 1 was merely one of convenience.

*L. parva incertae sedis* was isolated as a contaminant in EMJH medium (Hovind-Hougen et al., 1981). Morphologically it is similar to the leptospires, has phenotypic characteristics of both the pathogenic and the saprophytic forms, low DNA:DNA similarity values compared to other *Leptospira* sp. (Yasuda et al., 1987) and comparatively large evolutionary distance values (Table 2). Collectively these data support the ad hoc proposal for a new genus, *Turneria* (‘*L*. parva reference strain H; Anon., 1992).

*Leptonema illini* was unexpectedly more related to the *Leptospiraceae* than was *Leptospira parva incertae sedis* and the sequence data gave qualified support to it being a distinct genus (Hovind-Hougen, 1979).

The pathogenic leptospires (Table 1) have been divided broadly into two groups (*L. interrogans* and *L. kirschneri*, and *L. weilii*, *L. borgpetersenii* and *L. santarosai*) on the basis of such criteria as G + C content, sensitivity to 2,6-diaminopurine, lipase production (Haapala et al., 1969), DNA homologies (Ramadass et al., 1992) and the presence or absence of DNA fragments hybridizing on Southern blots to the sphA sphingomyelinase gene (Segers et al., 1992). These results are in agreement with those presented here though a third group, *L. noguchii* and *L. meyeri*, may be delineated (Fig. 1). Although the relatively low percentage bootstrap value indicated branching instability (Fig. 1), the *L. noguchii* and *L. meyeri* group shared a 16S rDNA sequence value of approximately 96.6%, an evolutionary distance value of 0.0154 (Table 2) and a mol% G + C of about 35.0% (Yasuda et al., 1987). Sequence values of between 95 and 97% are generally expected for different species within a genus (Paster & Dewhirst, 1988; Dewhirst et al., 1989).

The strain *L. meyeri ranarum ICF* was isolated from the kidneys of a leopard frog (*Rana pipiens*) and had the characteristics of being a saprophyte, growing at 13 °C and being unable to establish leptospirosis in laboratory animals (Diesch et al., 1966). Yet it formed part of a pathogenic 'complex', being sensitive to 8-azaguanine,
2,6-diaminopurine and copper sulphate (Brendle et al., 1974; Diesch et al., 1966; Ellinghausen, 1968). Furthermore, it was positive in a PCR test for pathogenic leptospires (Hookey, 1992), gave multiple putative sphingomyelinase gene patterns similar to strains of a pathogenic nature (Segers et al., 1992), an EcoRI 0.75-0.8 kb 16S rDNA fragment resembling the pathogens on 'ribotyping' (Hookey, 1993) and it occupied a central phylogenetic position within the pathogenic domain. In view of the weight of evidence, L. meyeri should be regarded as being related to the pathogens.

The type strain of the genus L. interrogans icterhaemorrhagiae RGA and L. kirschneri had about 99.84% sequence similarity and could be regarded as belonging to the same species. However, DNA:DNA hybridization values (approximately 31.1%) have clearly established both as separate entities (Yasuda et al., 1987; Ramadass et al., 1992). Similar results were also in evidence between the saprophytes, L. weilii and L. biflexa sensu stricto (Fig. 1; Table 2; Yasuda et al., 1987). The lack of correlation between 16S rDNA sequence similarity and DNA hybridization data is not unknown (Ash et al., 1991; Boddinghaus et al., 1990; Collins et al., 1989; Whitey et al., 1990) and it has been suggested that such recently diverged strains should be treated as part of a single rRNA species or super-species complex (Fox et al., 1992). In contrast, the serovar canicola may be regarded as a member of L. interrogans sensu stricto on both DNA relatedness (70%, Ramadass et al., 1992) and 16S rDNA sequence identity (Table 2).

The phylogenetic position of L. inadai was such that it bisected the pathogenic and saprophytic domains. Whether this represents an emerging subline or a new species, it would be a proposal to assign L. biflexa sensu lato ancona Ancona Porto to a new species.

In the absence of hybridization data and the relatively low consensus among the bootstrapped 16S rDNA sequence data sets (Fig. 1), little can be inferred of the genetic relationships between the freshwater isolates, L. biflexa sensu lato canela Canela and L. biflexa sensu lato jequitaia Jequitaia other than that they were distinct from each other and from the other strains examined.

In this study a comparative 16S rDNA sequencing analysis of all recognized species of Leptospiraceae revealed that the pathogenic species, though distinct from each other, comprised three groups and that the saprophytes, as currently circumscribed by L. biflexa sensu stricto and L. wolbachii, comprised a heterogeneous complex requiring further analyses. The lineages for L. inadai and L. parva incertae sedis were distinct. There was a correlation between groupings formed on total DNA hybridizations and those on 16S rDNA sequence data that may provide the framework for a natural classification of these spirochaetes.

The authors would like to thank Drs C. J. Duggleby, and N. A. Saunders for critical reading of this manuscript.

References


Dewhirst, F. E., Paster, B. J. & Bright, P. L. (1989). Chromobacterium, Eikenella, Kingella, Neisseria, Simonsiella, and Vitreoscilla species comprise a major branch of the beta group. Protobacteria by 16s ribosomal ribonucleic acid sequence comparison: type of Eikenella and Simonsiella the family.