Phylogenetic Analysis of the Genus *Borrelia*: a Comparison of North American and European Isolates of *Borrelia burgdorferi*

RICHARD T. MARCONI* AND CLAUDE F. GARON

Laboratory of Vectors and Pathogens, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, Hamilton, Montana 59840

Received 12 August 1991/Accepted 25 October 1991

We have sequenced the 16S rRNA molecules from four species of *Borrelia* and from six isolates of *Borrelia burgdorferi* via the reverse transcriptase primer extension method. The sequences were aligned and evolutionary relationships were determined, including the calculation of evolutionary distances and the construction of a phylogenetic tree. These analyses demonstrate significant divergence among *B. burgdorferi* isolates, with the European isolates G1 and G2 residing most distant from the main cluster. Signature nucleotides which distinguish *B. burgdorferi* from all other members of this genus and which distinguish the European isolates G1 and G2 from the North American isolates B31, Sh-2-82, and 1352 were identified. Finally, Southern blot analyses were performed to compare the restriction patterns of the genes coding for rRNA and to relate our data to the grouping scheme of Postic et al. (D. Postic, C. Edlinger, C. Richaud, F. Grimont, J. Dufresne, P. Perolat, G. Baranton, and P. A. D. Grimont, Res. Microbiol. 141:465-475, 1990).

Since the first cases of Lyme disease were diagnosed in Connecticut in 1975 (19), the number of diagnosed cases has increased each year. Lyme borreliosis is now the most frequently diagnosed arthropod-borne disease in the United States (10, 12). Since 1984, when *Borrelia burgdorferi* was first isolated, cultivated, and identified as the etiologic agent of Lyme disease (3), numerous isolates have been collected and studied.

Comparisons of different *B. burgdorferi* isolates by a variety of techniques and approaches have revealed considerable heterogeneity. Restriction analysis of whole-cell DNA has demonstrated unique cleavage patterns which appear to be isolate specific (10). Plasmid profiles have been found to differ from one isolate to another and to change with laboratory passage (17). By using the polymerase chain reaction, differences in the reactivity of various isolates with several different chromosomally derived primers have been observed (16). Differing restriction patterns of the genes coding for rRNA (rDNA) have recently been observed with Southern blot analysis (15). At the protein level, there appear to be significant differences in the reactivities of particular isolates with monoclonal antibodies (4, 22).

The clinical manifestations of Lyme borreliosis in Europe and in North America appear to differ as well (18). Acrodermatitis chronica atrophicans, a severe dermatologic disorder, and neuroborreliosis are common in chronically infected patients in Europe yet relatively uncommon in the United States. The converse appears to be true for the arthritic manifestations, which are common in the United States yet relatively uncommon in Europe (1, 2, 18). While the underlying explanation(s) for differences in the clinical manifestations remains unclear, it is conceivable that genetic differences among the infecting organisms may be responsible.

To assess the genetic divergence among borrelial isolates, including North American and European isolates of *B. burgdorferi*, we have conducted a phylogenetic analysis based on their 16S rRNA sequences. Included in this anal-

ysis are *B. hermsii*, *B. coriaceae*, and *B. anserina*, North American isolates of *B. burgdorferi* (B31, Sh-2-82, and 1352), and European isolates of *B. burgdorferi* (G1, G2, and 20004).

The use of rRNA sequences as indicators of evolutionary divergence is now widespread, and its advantage over classical taxonomic approaches is well documented (23). To date, more than 500 small subunit rRNA sequences have been determined and used for this purpose (13). Here we present in near entirety the 16S rRNA sequences of several borrelial species and isolates. The phylogenetic relationships among the isolates were determined through calculations of percent sequence similarity and evolutionary distances (defined as the estimated number of nucleotide substitutions per sequence position) (6) and presented as a phylogenetic tree. In addition, rDNA restriction patterns were determined with Southern blot analyses, and the data were compared with the genomic grouping scheme of Postic et al. (15).

MATERIALS AND METHODS

Bacterial isolates and growth. *Borrelia* isolates employed in this study were kindly provided by Tom Schwan of the Rocky Mountain Laboratories. *B. burgdorferi* B31 (ATCC 35210) and Sh-2-82 originated from *Ixodes dammini* ticks, Shelter Island, N.Y.; *B. burgdorferi* G1 and G2 were cerebrospinal fluid isolates from Lyme patients in Germany; *B. burgdorferi* 1352 originated from the tick *Amblyomma americanum* from Texas; *B. burgdorferi* 20004 originated from *Ixodes ricinus* from France; *B. hermsii* serotype C (ATCC 35209) originated from *Ornithodoros hermsii* in California; *B. coriaceae* Co53 (ATCC 43381) originated from *Ornithodoros coriaceus* from California; and *B. anserina* was from the Rocky Mountain Laboratories bacterial pathogen collection. Cells were cultured in BSK-II medium (3) at 35°C.

RNA isolation and sequencing. Total cellular RNA was isolated from 500-ml cultures as previously described (21). The reverse transcriptase dideoxynucleotide primer extension method used was derived from that described by Tappich et al. (20). Modifications included the deletion of

* Corresponding author.

FIG. 1. Phylogenetic tree of the genus *Borrelia* and *B. burgdorferi* isolates. The tree was constructed as described in the text. Circled numbers indicate the number of times out of 100 that a particular node was supported by bootstrap analysis. Evolutionary distances are measured along line segments. The bar represents an evolutionary distance, as defined in the text, of 0.003.

the chase reaction and the use of end-labeled primers. Primers were synthesized with an automated DNA synthesizer and were 5' end labeled as previously described (5). Sequencing gels were prepared and run by standard techniques.

DNA isolations, restriction digests, and Southern blot analysis. Total *Borrelia* DNA was isolated as described by Barbour et al. (4). Restriction enzymes were purchased from Boehringer Mannheim, and 7 μ g of whole-cell DNA was digested by standard methods. Digested DNA was loaded directly onto 0.7% agarose gels (10 by 15 cm) in TAE (40 mM Tris-acetate–2 mM EDTA, pH 8.5) and electrophoresed at 25 V for 20 h. DNA was vacuum blotted onto GeneScreen membranes (DuPont-NEN) with the VacuGene system (Pharmacia LKB). Hybridizations were carried out as described in the GeneScreen manual. Blots were probed with nick-translated plasmid pKK3535. This plasmid carries the *Escherichia coli* *rrnB* operon. The plasmid was labeled by using a nick translation kit (Boehringer Mannheim) and [α - 32 P]ATP (3,000 Ci/mmol; NEN).

Phylogenetic analysis. The nine *Borrelia* 16S rRNA sequences were aligned by using the Clustal alignment program of the PCGENE software package (Intellegentics, Inc.). Masked sequences of 1,347 bases were generated by removing gaps and unidentified nucleotide positions. Evolutionary analyses were performed with the PHYLIP program package written by Joseph Felsenstein (6). Evolutionary distances were obtained with the Jukes-Cantor option of the DNADIST program. By using these distance values, a phylogenetic tree was constructed with the FITCH program (7), with the 16S rRNA sequence from the *rrnB* operon of *E. coli* serving as an out-group. Tree topology was reexamined by using bootstrapped datum sets (100 sets) obtained with the DNABOOT program, which was followed by parsimony analysis with the DNAPARS program (both programs are contained in the PHYLIP package). The confidence levels for each node are indicated in Fig. 1.

Nucleotide sequence accession numbers. The 16S rRNA sequences have been entered in the EMBL/GenBank data bases with the following accession numbers: *B. burgdorferi* isolates B31, Sh-2-82, 1352, 20004, G1, and G2 have accession numbers X57404, M60969, M64309, M64311, and

M60967, respectively, and *B. hermsii*, *B. coriaceae*, and *B. anserina* have accession numbers M60968, M60970, and M64312, respectively.

RESULTS AND DISCUSSION

We have analyzed the genetic relationships between North American and European isolates of *B. burgdorferi* as well as among several species of *Borrelia* through the determination and comparison of their 16S rRNA sequences. The initial sequencing primer complemented a 15-base highly conserved sequence of *E. coli* 16S rRNA spanning bases 1492 to 1506 (13). On the basis of the sequence obtained, additional primers were synthesized and utilized in the sequencing reactions. A total of 16 different primers were used to sequence each 16S rRNA so as to result in significant overlap of the determined sequences. The sequences presented here are partial and span approximately 95% of the 16S rRNA molecule. Masked sequences were generated and used to calculate evolutionary distances and percent sequence similarity values (Table 1) and to construct a phylogenetic tree (Fig. 1).

16S rRNA sequences were determined for four species of the *Borrelia* genus, including *B. anserina* (the type species for the genus), *B. hermsii*, *B. coriaceae*, and *B. burgdorferi*. *B. hermsii* and *B. coriaceae* were found to be the two most closely related species of the genus, with a sequence similarity value of 98.9%. *B. coriaceae* is the most recent member of the *Borrelia* genus to be identified. The classification of this organism as a distinct species, which has been inferred from restriction endonuclease analysis and DNA hybridization studies (9, 10), is supported by the evolutionary distances and phylogenetic relationships inferred from the analysis of 16S rRNA sequences. *B. hermsii*, *B. coriaceae*, and *B. anserina* as a group can be distinguished from the *B. burgdorferi* isolates sequenced here on the basis of signature or diagnostic nucleotide positions. The locations and identities of these residues are listed in Table 2. It is interesting to note that this division also reflects the relationships of the *Borrelia* spp. with their arthropod vectors. *B. hermsii*, *B. anserina*, and *B. coriaceae* all share the characteristic of being carried by soft-bodied ticks. *B. burgdorferi*

TABLE 1. Percent sequence similarity values

Isolate or species	% Sequence similarity									
	B31	Sh-2-82	1352	G1	G2	20004	<i>B. hermsii</i>	<i>B. coriaceae</i>	<i>B. anserina</i>	<i>E. coli</i>
B31		99.9	99.9	99.1	99.3	99.9	96.7	96.6	96.4	71.1
Sh-2-82			99.8	99.0	99.2	99.2	96.7	96.5	96.3	71.0
1352				99.1	99.2	99.9	96.7	96.6	96.3	70.9
G1					99.8	99.2	96.3	96.5	95.9	70.9
G2						99.3	96.4	96.7	96.0	70.9
20004							96.7	96.7	96.4	71.0
<i>B. hermsii</i>								98.9	98.4	70.3
<i>B. coriaceae</i>									99.2	70.6
<i>B. anserina</i>										70.9
<i>E. coli</i>										

is distinct from this group in that it is carried by hard-bodied ticks. *B. tyleri* represents the only other species of *Borrelia* known to be carried by a hard-bodied tick. The base signature identity of this organism is not currently available.

In order to evaluate the degree of genetic divergence among isolates of *B. burgdorferi*, three North American (B31, Sh-2-82, and 1352) and three European (G1, G2, and 20004) isolates were analyzed. The North American isolates, B31, 1352, and Sh-2-82, were found to be tightly clustered and very closely related. In contrast, considerable divergence among the European isolates was observed. The French isolate, 20004, was found to be more closely related to the North American isolates than to the European isolates. The heterogeneity of the European isolates is also evident from signature nucleotide analysis. Several signature positions distinguish the European isolates G1 and G2, but not 20004, from their North American counterparts. These signatures occurred at positions 91, 92, 140, 183, 187, 278, 630, 649, and 979 (*E. coli* nomenclature).

The phylogenetic tree presented in Fig. 1 was constructed with *E. coli* serving as the out-group. In order to further verify the observed branching patterns, bootstrap analysis was performed. All of the main branch points, i.e., those between established species and that between the two clusters of *B. burgdorferi* isolates, were fully supported, occurring in 100% of the bootstrap replicates. These results provide strong support for the classification of these isolates into two distinct genomic groups.

We have also employed a Southern blot analysis of rDNA restriction patterns to assess the genetic relationships among the *Borrelia* spp. rDNA restriction pattern analysis has been

used as a taxonomic or epidemiological tool in the investigation of the *Leptospira* spp. (14), the *Legionella* spp. (8), and *B. burgdorferi* (15). Postic and coworkers (15) compared the DNA relatedness, via hybridization studies, and the rDNA restriction profiles of 16 different isolates of *B. burgdorferi* from Europe and North America. On the basis of those results, the authors suggested that *B. burgdorferi* isolates are represented by two genomic species which they referred to as the *B. burgdorferi* sensu stricto group and the 20047 group. The rDNA restriction patterns obtained after digestion with *EcoRV* were assessed via Southern blot analysis (Fig. 2). Analysis of *EcoRV*-digested DNA revealed two different patterns. The European isolate 20004 and the North American isolates were characterized by two bands of approximately 6.5 and 4.1 kb. The European isolates G1 and G2 exhibited a profile composed of three bands at approximately 5, 4.1, and 1.5 kb. While the band at 1.5 kb is consistently observed, visualization often requires prolonged exposure of the film. By the classification scheme proposed by Postic et al. (15), isolates Sh-2-82, 1352, B31, and 20004 would be placed in the *B. burgdorferi* sensu stricto group, while isolates G1 and G2 would reside in the 20047 group. We should also note that the patterns observed here

TABLE 2. *B. burgdorferi* signature nucleotide positions

Position(s) ^a	Nucleotide ^b	
	<i>B. burgdorferi</i>	<i>B. hermsii</i> , <i>B. coriaceae</i> , and <i>B. anserina</i>
65, 433, 593, 604, 645, 665, 842, 1,006, 1,040, 1,120, 1,154	A	G
130, 721, 1,117, 1,140, 1,304	G	A
650, 193, 1,138	U	A
408, 594, 636, 646, 679, 840, 1,001, 1,172, 1,471	U	C
192	U	G
840	C	U

^a The *E. coli* numbering scheme is used.

^b Standard nucleotide abbreviations are used.

FIG. 2. Southern blot analyses of *EcoRV* rDNA restriction patterns. All methods were as described in the text.

for B31 and *B. hermsii* are consistent with those reported by Postic et al. (15). The genetic grouping inferred from this approach is consistent with that obtained via 16S rRNA sequence analysis.

In this report, we have analyzed the genetic relationships among isolates of this medically important genus on the basis of 16S rRNA sequences and Southern blot analyses of rDNA restriction patterns. The analyses have quantified the genetic differences among various North American and European isolates of *B. burgdorferi*. On the basis of the results presented here and the previous work mentioned above, the designation of each of these genomic groups as a distinct species, or at least subspecies, warrants consideration. In addition, the identification of signature positions which define these organisms, both at the species and at the isolate level, has allowed for the development of diagnostic probes directed at these sites (11). With the determination of additional 16S rRNA sequences from other isolates and species, we will gain significantly in our understanding of the basic biology and genetic relationships among the *Borrelia* spp. and perhaps draw closer to an understanding of the factors which result in different clinical manifestations of Lyme disease in different geographic regions.

ACKNOWLEDGMENTS

We thank T. Schwan, R. Blakemore, W. Cieplak, R. Gilmore, and D. Dorward for their helpful comments and critique in the preparation of the manuscript. We thank R. Evans and G. Hettrick for graphics and photography and B. Kester for typing the manuscript.

REFERENCES

- Ackermann, R., B. Rehse-Kupper, E. Gollmer, and R. Schmidt. 1988. Chronic neurologic manifestations of erythema migrans borreliosis. *Ann. N. Y. Acad. Sci.* **539**:16-23.
- Asbrink, E., and A. Hovmark. 1988. Early and late cutaneous manifestations in *Ixodes*-borne borreliosis (erythema migrans borreliosis, Lyme borreliosis), p. 4-15. In J. L. Benach and E. M. Bosler (ed.), *Lyme disease and related disorders*. New York Academy of Sciences, New York.
- Barbour, A. G. 1984. Isolation and cultivation of Lyme disease spirochetes. *Yale J. Biol. Med.* **57**:521-525.
- Barbour, A. G., R. A. Heiland, and T. R. Howe. 1985. Heterogeneity of major proteins in Lyme disease borreliae: a molecular analysis of North American and European isolates. *J. Infect. Dis.* **152**:478-484.
- Chaconas, G., and J. H. van De Sande. 1980. $5'$ - 32 P labeling of RNA and DNA restriction fragments. *Methods Enzymol.* **65**:75-88.
- Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* **57**:379-404.
- Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. *Science* **155**:278-284.
- Grimont, F., M. Lefevre, E. Ageron, and P. A. D. Grimont. 1989. rRNA gene restriction patterns of *Legionella* species: a molecular identification system. *Res. Microbiol.* **140**:615-626.
- Johnson, R. C., W. Burgdorfer, R. S. Lane, A. G. Barbour, S. F. Hayes, and F. W. Hyde. 1987. *Borrelia coriaceae* sp. nov.: putative agent of epizootic bovine abortion. *Int. J. Syst. Bacteriol.* **37**:454-456.
- LeFebvre, R. B., G. C. Perng, and R. C. Johnson. 1989. Characterization of *Borrelia burgdorferi* isolates by restriction endonuclease analysis and DNA hybridization. *J. Clin. Microbiol.* **27**:636-639.
- Marconi, R. T., L. Lubke, W. Hauglum, and C. F. Garon. Submitted for publication.
- Miller, G. L., R. B. Craven, R. E. Bailey, and T. F. Tsai. 1990. The epidemiology of Lyme disease in the United States 1987-1988. *Lab. Med.* **21**:285-289.
- Neefs, J. M., Y. Van de Peer, L. Hendriks, and R. Wachter. 1990. Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res.* **18**:2237-2317.
- Perolat, R., F. Grimont, B. Gernault, P. A. D. Grimont, E. Fournie, H. Thevenet, and G. Baranton. 1990. rRNA gene restriction patterns of *Leptospira*; a molecular typing system. *Res. Microbiol.* **141**:159-171.
- Postic, D., C. Edlinger, C. Richaud, F. Grimont, J. Dufresne, P. Perolat, G. Baranton, and P. A. D. Grimont. 1990. Two genomic species in *Borrelia burgdorferi*. *Res. Microbiol.* **141**:465-475.
- Rosa, P. A., D. Hogan, and T. G. Schwan. 1991. Polymerase chain reaction analyses identify two distinct classes of *Borrelia burgdorferi*. *J. Clin. Microbiol.* **29**:524-532.
- Schwan, T. G., W. Burgdorfer, and C. F. Garon. 1988. Changes in infectivity and plasmid profile of the Lyme disease spirochete, *Borrelia burgdorferi*, as a result of in vitro cultivation. *Infect. Immun.* **56**:1831-1836.
- Steere, A. C. 1989. Lyme disease. *N. Engl. J. Med.* **321**:586-596.
- Steere, A. C., S. E. Malawista, D. R. Snyderman, R. E. Shope, W. A. Andiman, M. R. Ross, and F. M. Steele. 1977. Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. *Arthritis Rheum.* **20**:7-17.
- Tapprich, W. E., D. J. Goss, and A. E. Dahlberg. 1989. Mutation at position 791 in *Escherichia coli* 16S ribosomal RNA affects processes involved in the initiation of protein synthesis. *Proc. Natl. Acad. Sci. USA* **86**:4927-4931.
- Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* **77**:5201-5205.
- Wilske, B., V. Preac-Mursic, G. Schierz, and K. V. Busch. 1986. Immunochemical and immunological analysis of European *Borrelia burgdorferi* strains. *Zentralbl. Bakteriol. Mikrobiol. Hyg. Ser. A* **263**:92-102.
- Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221-271.