Identification of a third genomic group of *Borrelia burgdorferi* through signature nucleotide analysis and 16S rRNA sequence determination

RICHARD T. MARCONI* and CLAUDE F. GARON

Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Allergy and Infectious Diseases, Laboratory of Vectors and Pathogens, Rocky Mountain Laboratories, Hamilton, Montana 59840, USA

(Received 9 September 1991; revised 18 November 1991; accepted 3 December 1991)

As part of a continuing effort to assess genetic variation among isolates of *Borrelia burgdorferi* we have determined the 16S rRNA signature nucleotide makeup of two tick isolates from the USSR. Signature nucleotides were identified via reverse transcriptase primer extension sequencing of select regions of the 16S rRNA molecule. In addition, the near complete 16S rRNA sequence of one of the isolates, R-IP3, was determined and utilized in a phylogenetic assessment. The sequence was aligned with the 16S rRNA sequences of other *B. burgdorferi* isolates as well as with other *Borrelia* species. Distance matrix analyses were performed and a phylogenetic tree was constructed. These analyses demonstrate that these isolates belong to a third previously unidentified genomic group of *B. burgdorferi*.

Introduction

The causative agent of Lyme borreliosis, *Borrelia burgdorferi*, was first isolated from *Ixodes dammini* ticks on Shelter Island, NY, USA, in 1982 (Burgdorfer *et al.*, 1982). Isolates have now been collected worldwide and characterized by several approaches including rRNA gene restriction profiles (Postic *et al.*, 1990; Marconi & Garon, 1992), reactivity with various PCR primers (Rosa *et al.*, 1991; Malloy *et al.*, 1990), reactivity with monoclonal antibodies (Wilske *et al.*, 1985, 1988) and via phylogenetic analysis of 16S rRNA sequences (Marconi & Garon, 1992). These studies suggest that isolates of *B. burgdorferi* can be typed into at least two genomic classes now referred to as the *sensu stricto* and 20047 groups (Postic *et al.*, 1990). Since the clinical presentations of Lyme disease have been found to differ among geographic regions (Steere, 1989), an understanding of the genetic divergence within this species may provide information to help explain the clinical variation.

In this study, we have investigated the phylogenetic relationship of two Russian isolates of *B. burgdorferi* to isolates from other geographic regions. Based upon previously determined *Borrelia* 16S rRNA sequences (Marconi & Garon, 1992), we have constructed oligonucleotide primers which allow for the rapid sequence

determination of select regions of the 16S rRNA molecule which contain signature nucleotides specific for each genomic group (Marconi *et al.*, 1992). Comparison of the full complement of signature nucleotides suggests that they belong to a previously unidentified genomic group. To further characterize the relationship of these isolates with others, we determined the near complete 16S rRNA sequence for isolate R-IP3, and conducted phylogenetic analyses. Distance matrix analyses were performed and a phylogenetic tree was constructed. These results support the designation of an additional genomic group of *B. burgdorferi*.

Methods

Bacterial strains and culture conditions. *Borrelia* isolates and species used in this study were provided by Dr Tom Schwan (this laboratory). Isolates R-IP3 and R-IP21 were originally isolated by Kryuchnikov *et al.* from *Ixodes persulcatus* ticks from the Leningrad region of the USSR (Kryuchnikov *et al.*, 1988).

RNA isolation, sequence determination, and phylogenetic analysis. All methods were as described by Marconi & Garon (1992). For selective sequencing through regions containing diagnostic signature nucleotide positions, four different primers designated as primers 1000 (5'-GGGGAATAATTATCTCTAAC), 690 (5'-ATCAACAGATTG-CACCCTTAC), 400 (5'-CCCATTGCGGAAGATTCTTAG) and 190 (5'-TTAAAGGCTTCCTTCATC) were used. Primers were 5'-end-labelled as described by Chaconas & van De Sande (1980).

For phylogenetic analysis, the near complete sequence of isolate R-IP3 was determined and from this a masked sequence of 1347 bases

* Author for correspondence. Tel. (406) 363 3211; fax (406) 363 6406.

was generated by removing gap and unidentified nucleotide positions. The 16S rRNA sequence for isolate R-IP3 has been assigned the GenBank accession number, M75149. The origin of the isolates used for comparative purposes and their 16S rRNA sequences have been described previously (Marconi & Garon, 1992). Evolutionary analyses were performed as previously described (Marconi & Garon, 1992) with the PHYLIP program package written by Joseph Felsenstein (Felsenstein, 1982).

Results and Discussion

The comparison of 16S rRNA sequences is now widely used as a tool in assessing the genetic relationships among organisms (for a review, see Woese, 1987). The rRNA database currently contains in excess of 500 16S rRNA sequences. Thus, at the sequence level, 16S rRNA represents perhaps the most intensively studied cellular molecule. Earlier phylogenetic studies demonstrated that upon alignment of the 16S rRNA sequences, signature nucleotides could be identified which were kingdom-specific (Woese, 1987). It is now recognized that signature positions can be defined at the species and subspecies level. These signature nucleotides have been exploited in nucleic-acid-based identification systems, directed against specific rRNA sequences, which are species- or subspecies-specific. Nucleic-acid-based detection systems have been developed for *Mycoplasma* (Göbel *et al.*, 1987), *Treponema* (Jensen *et al.*, 1990), *Campylobacter* (Cox *et al.*, 1990), *Haemophilus* (Parsons *et al.*, 1989), *Borrelia* (Marconi *et al.*, 1992) and *Clostridium* (Wilson *et al.*, 1988).

Currently, two genomic groups of *B. burgdorferi* are recognized. Postic and co-workers provided the first definition of these genomic groups based primarily on whole-cell DNA/DNA hybridization and also upon RFLP patterns of the rRNA genes. The two genomic groups were termed the *sensu stricto* and 20047 groups (Postic *et al.*, 1990). Previously, we assessed the genetic variation among isolates of *B. burgdorferi* through RFLP mapping of the rRNA genes and via a phylogenetic analysis based on 16S rRNA sequences (Marconi & Garon, 1992). The results of that study were consistent with those of Postic *et al.* (1990) and demonstrated the existence of at least two genomic groups of *B. burgdorferi*. In addition, signature nucleotides were defined which distinguish between the genomic groups.

In this report, we have utilized a sequencing strategy which allows for determination of the genetic classification of *B. burgdorferi* isolates based upon their 16S rRNA signature nucleotide makeup. The strategy involves the use of extension primers which allow for the determination of all the genomic-group-specific signature nucleotides. Thus, the genomic group can be determined without complete sequencing of the 16S

Table 1. Comparison of the signature nucleotides of *B. burgdorferi* isolates

Position*	Genomic group		Isolate	
	<i>Sensu stricto</i> †	20047‡	R-IP3	R-IP21
91	U	C	U	U
92	C	U	U	U
140	U	C	U	U
183	G	A	G	G
187	G	A	A	A
278	A	G	G	G
630	C	A	G	G
649	G	A	G	G
979	C	U	C	C

* Positions are numbered according to *E. coli* nomenclature.

† Isolates belonging to this genomic group for which 16S rRNA sequences have been determined and their GenBank accession numbers include B31 (X57404), Sh-2-82 (M60969), 1352 (M64309) and 20004 (M64310) (Marconi & Garon, 1992).

‡ Isolates belonging to this group for which 16S rRNA sequences have been determined include G1 (M64311) and G2 (M60967) (Marconi & Garon, 1992).

rRNA molecule. We have determined the signature composition of two tick (*I. persulcatus*) isolates, R-IP3 and R-IP21 from the USSR. Sequence analysis revealed that both of the Russian isolates possess identical signature identities. The composition was found to be distinct from that which we had previously identified for members of the *sensu stricto* and 20047 genomic groups. Table 1 compares the signature composition of these isolates with those of the established genomic groups. It is important to point out that isolate R-IP3, which was isolated by Kryuchnikov *et al.* (1988) and investigated in this study, is distinct from isolate IP3 described by Postic *et al.* (1990). Isolate IP3, a cerebrospinal fluid isolate from France, was determined to be a member of the *sensu stricto* genomic group by those authors (Postic *et al.*, 1990). Recently it was proposed that the *sensu stricto* genomic group could be further subdivided into at least three subgenomic groups by Adam *et al.* (1991). These authors used a variety of approaches including sequence analysis of a limited portion of the 16S rRNA molecule of three members of the *sensu stricto* group, B31, pKO and IP3, in reaching this conclusion. The subgroups were defined in part, based upon a one nucleotide difference in their 16S rRNA sequences. They refer to these variable positions as signature positions. In that there are nine genomic-group-specific nucleotide positions (refer to Table 1), the significance of subdividing genomic groups based upon a difference at one such position is unclear. The designation of additional unique genomic groups, based upon minor differences in the set of signature nucleotides, should only be considered when

Fig. 1. Phylogenetic tree of selected *Borrelia* species and *B. burgdorferi* isolates. Phylogenetic analysis and tree construction were accomplished through use of the PHYLIP (version 3.4) package as described in Methods. The scale bar indicates a percentage sequence similarity difference of 0.150%. The numbers enclosed in circles at each branch node indicate the number of times out of 100 that a particular node was supported by bootstrap analysis.

those isolates are also demonstrated to exhibit a significant degree of evolutionary distance.

In order to quantitatively assess the phylogenetic divergence of these isolates from members of the other groups, the near complete 16S rRNA sequence of isolate R-IP3 was determined. Through the reverse transcriptase primer extension method, a 1475 nucleotide segment was determined, representing approximately 95% of the molecule. Sixteen different extension primers were used for sequencing. This resulted in significant overlap in the determined sequences. Via this approach, sequences could be verified, thus addressing concerns about the error frequency of reverse transcriptase sequencing.

The 16S rRNA sequence from isolate R-IP3 was aligned with those from several other *B. burgdorferi* isolates and *Borrelia* species (data not shown). The isolates, isolation source, and their geographic origin have been described previously (Marconi & Garon, 1992). From the aligned sequences, masked sequences were generated by removing positions corresponding to gaps or ambiguous residues. The aligned masked sequences were subjected to distance matrix analyses using the DNADIST program as described in Methods. When compared to the type strain B31, isolate R-IP3 represents the most distantly related member of the species characterized to date with a percentage sequence similarity value of 98.9%. The percentage sequence similarity values are presented in Table 2.

In order to determine the phylogenetic branching patterns, a phylogenetic tree was constructed (Fig. 1). The *Escherichia coli* 16S rRNA sequence served as the outgroup. The branching pattern demonstrates that isolate R-IP3 resides peripherally to both the *sensu stricto* and 20047 genomic groups. The branching pattern observed was re-evaluated through bootstrap analysis

Table 2. R-IP3 16S rRNA percentage sequence similarity values

<i>B. burgdorferi</i> isolate			<i>Borrelia</i> spp.				
B31	1352	G1	<i>B. anserina</i>	<i>B. coriaceae</i>	<i>B. hermsii</i>	<i>E. coli</i>	
98.9	99.0	98.8	96.5	96.5	96.7	70.7	

using the DNABOOT program (also contained within the PHYLIP package). One hundred subreplicates were analysed. Bootstrap analysis demonstrated that the branching nodes could be accepted with a high degree of confidence. We have previously noted that isolates which share identical signature nucleotide makeups are branched from the same node (Marconi & Garon, 1992). Thus, while the complete sequence of isolate R-IP21 was not determined, it is reasonable to assume it would also branch with the R-IP3 isolate.

The results presented here warrant the designation of a third genomic group of *B. burgdorferi*. We propose that this group be referred to as genomic group 3. As we reported earlier for the 20047 genomic group (Marconi & Garon, 1992), the evolutionary distances which separate members of genomic group 3 from the other groups are of sufficient magnitude that the consideration of this group as a distinct species of *Borrelia* warrants further consideration. In the course of determining the signature identities of these and other isolates we have noted that a correlation may exist between signature makeup and the species of the arthropod vector (unpublished results, R. T. Marconi). Establishment of such a correlation will require the determination of the signature composition of several additional tick isolates of *B. burgdorferi*. If a

correlation can be clearly established then knowledge of the signature composition of an isolate may provide useful epidemiological information. We are continuing to address this question.

References

- ADAM, T., GASSMANN, G. S., RASIAH, C. & GÖBEL, U. B. (1991). Phenotypic and genotypic analysis of *Borrelia burgdorferi* isolates from various sources. *Infection and Immunity* **59**, 2579–2585.
- BURGDORFER, W., BARBOUR, A. G., HAYES, S. F., BENACH, J. L., GRUNWALDT, E. & DAVIS, J. P. (1982). Lyme disease – a tick-borne spirochetosis. *Science* **216**, 1317–1319.
- CHACONAS, G. & VAN DE SANDE, J. H. (1980). $5'$ - ^{32}P labelling of RNA and DNA restriction fragments. *Methods in Enzymology* **65**, 75–88.
- COX, N., JOHNSTON, J., SZARKA, Z., WRIGHT, D. J. M. & ARCHARD, L. C. (1990). Characterization of an rRNA gene-specific cDNA probe: applications in bacterial identification. *Journal of General Microbiology* **136**, 1639–1643.
- FELSENSTEIN, J. (1982). Numerical methods for inferring evolutionary trees. *Quarterly Review of Biology* **57**, 379–404.
- GÖBEL, U. B., GEISER, A. & STANBRIDGE, E. J. (1987). Oligonucleotide probes complementary to variable regions of ribosomal RNA discriminate between *Mycoplasma* species. *Journal of General Microbiology* **133**, 1969–1974.
- JENSEN, N. S., CASEY, T. A. & STANTON, T. B. (1990). Detection and identification of *Treponema hyodysenteriae* by using oligodeoxynucleotide probes complementary to 16S rRNA. *Journal of Clinical Microbiology* **28**, 2717–2721.
- KRYUCHECHNIKOV, V. N., KORENBERG, E. I., SHCHERBAKOV, S. V., KOVALEVSKY, YU. V. & LEVIN, M. L. (1988). Identification of *Borrelia* isolated in the USSR from *Ixodes persulcatus* schulze ticks. *Journal of Microbiology, Epidemiology and Immunobiology* **12**, 41–44.
- MALLOY, D. C., NAUMAN, R. K. & PAXTON, H. (1990). Detection of *Borrelia burgdorferi* using the polymerase chain reaction. *Journal of Clinical Microbiology* **28**, 1089–1093.
- MARCONI, R. T. & GARON, C. F. (1992). Phylogenetic analysis of the genus *Borrelia*: a comparison of North American and European Isolates of *Borrelia burgdorferi*. *Journal of Bacteriology* **174**, 241–244.
- MARCONI, R. T., LUBKE, L., HAUGLUM, W. & GARON, C. F. (1992). Species-specific identification of and distinction between *Borrelia burgdorferi* genomic groups by using 16S rRNA-directed oligonucleotide probes. *Journal of Clinical Microbiology* (in the Press).
- PARSONS, L. M., SHAYEGANI, M., WARING, A. L. & BOPP, L. H. (1989). DNA probes for the identification of *Haemophilus ducreyi*. *Journal of Clinical Microbiology* **27**, 1441–1445.
- POSTIC, D., EDLINGER, C., RICHAUD, C., GRIMONT, F., DUFRESNE, J., PEROLAT, P., BARANTON, G. & GRIMONT, P. A. D. (1990). Two genomic species in *Borrelia burgdorferi*. *Research in Microbiology* **141**, 465–475.
- ROSA, P. A., HOGAN, D. & SCHWAN, T. G. (1991). Polymerase chain reaction analyses identify two distinct classes of *Borrelia burgdorferi*. *Journal of Clinical Microbiology* **29**, 524–532.
- SIMPSON, W. J., SCHRUMPF, M. E. & SCHWAN, T. G. (1990). Reactivity of human Lyme borreliosis sera with a 39-kilodalton antigen specific to *Borrelia burgdorferi*. *Journal of Clinical Microbiology* **28**, 1329–1337.
- STEERE, A. (1989). Lyme disease. *New England Journal of Medicine* **321**, 586–596.
- WILSKE, B., PREAC-MURSIC, V. & SCHIERZ, G. (1985). Antigenic heterogeneity of European *Borrelia burgdorferi* strains isolated from patients and ticks. *Lancet* **1**, 1099.
- WILSKE, B., PREAC-MURSIC, V., SCHIERZ, G., KUHBECK, R., BARBOUR, A. G. & KRAMER, M. (1988). Antigenic variability of *Borrelia burgdorferi*. *Annals of the New York Academy of Sciences* **539**, 126–143.
- WILSON, K. H., BLITCHINGTON, R., HINDENACH, B. & GREENE, R. C. (1988). Species-specific oligonucleotide probes for rRNA of *Clostridium difficile* and related species. *Journal of Clinical Microbiology* **26**, 2484–2488.
- WOESE, C. R. (1987). Bacterial evolution. *Microbiological Reviews* **51**, 221–271.