JOURNAL OF BACTERIOLOGY, Aug. 1994, p. 4572–4582
0021-9193/94/$04.00+0

Vol. 176, No. 15

Analysis of the Distribution and Molecular Heterogeneity of the *ospD* Gene among the Lyme Disease Spirochetes: Evidence for Lateral Gene Exchange

RICHARD T. MARCONI,* D. SCOTT SAMUELS, ROBERT K. LANDRY, AND CLAUDE F. GARON

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

Received 9 February 1994/Accepted 24 May 1994

Analysis of the *ospD* gene has revealed that this gene is not universal among Lyme disease spirochete isolates. The gene was found to be carried by 90, 50, and 24% of the *Borrelia garinii*, *B. afzelii*, and *B. burgdorferi* isolates tested. Size variability in the *ospD*-encoding plasmid was also observed. Sequence analysis has demonstrated the presence of various numbers of a 17-bp repeated sequence in the upstream control (promoter) region of the gene. In addition, a region within the coding sequence where various insertions, deletions, and direct repeats occur was identified. *ospD* gene sequences from 31 different isolates were determined and utilized in pairwise sequence comparisons and construction of a gene tree. These analyses suggest that the *ospD* gene was the target of several recombinational events and that the gene was recently acquired by Lyme disease spirochetes and laterally transferred between species.

Several recent studies have demonstrated that the causative agent of Lyme disease, the *Borrelia burgdorferi* sensu lato complex (10, 21, 38), comprises at least three distinct *Borrelia* species (2, 26, 42). A revised nomenclature has been proposed (2), designating these species *B. burgdorferi*, *B. garinii* (1a), and *B. afzelii* sp. nov. (11) (previously referred to as group VS461 [2] or group 3 [26]). Several of the outer surface proteins (Osps) associated with these Lyme disease spirochete (LDS) species, including OspA, OspB (7), and OspC (19), have been identified and characterized. Recently a fourth Osp, OspD, has been identified (32). All have been demonstrated to be lipoproteins (9, 32). OspA, OspB, and OspC have been the major focus in efforts to develop a Lyme disease vaccine (17, 18, 33) or diagnostic test (39). As an extension of these efforts, the variability of these proteins among the three LDS species has been intensively studied (22, 28, 34, 42, 43).

Borreliae are unique among eubacteria in that their genome is organized into a linear chromosome of approximately 960 kb (6, 12, 16, 30, 36) and a series of covalently closed linear and circular plasmids (4, 5). Interestingly, all of the genes encoding Osps have been mapped to plasmids. OspA and OspB are encoded by genes on linear plasmids of 50, 55, and 56 kb in *B. burgdorferi* (5), *B. garinii*, and *B. afzelii*, respectively (36). In contrast, *ospC* maps to a 26-kb circular plasmid and is the only gene to be mapped to a circular DNA molecule in any *Borrelia* species (29, 35). Norris et al. (32) have recently demonstrated that in *B. burgdorferi* B31, the *ospD* gene is carried on a 38-kb linear plasmid and that this plasmid can be lost upon prolonged in vitro cultivation. High-passage strain B31, lacking the *ospD*-encoding plasmid, had lost the ability to infect in animal models. On the basis of these observations, the researchers suggested that OspD may play a role in the establishment of infection or in virulence. If OspD does participate in these processes, then it may be of practical importance with regard to the diagnosis and control of Lyme disease.

In this study, we sought to assess the distribution of the *ospD* gene among the LDS species and to analyze its variability at the sequence level. In addition, we wished to determine if the repeat elements identified by Norris et al. (32) in the control region of *B. burgdorferi* B31 *ospD* are conserved among isolates. We found that while the gene can be found in at least some isolates of each of the three LDS species, it is not universal among isolates. Sequence analysis demonstrated that the level of conservation of this gene is greater than that observed for other characterized Osps (14, 15, 22, 40, 42, 43). However, the gene appears to contain several areas where a high degree of recombination has occurred. The results of pairwise sequence comparisons and gene tree construction demonstrate that the *ospD* sequences do not yield a gene tree consistent with those obtained by analysis of other LDS genes (13, 26). While most analyses delineate three phyletic groups, the *ospD*-derived gene tree delineates only two groups, with *B. garinii* and *B. afzelii* forming a single cluster. One possible interpretation of the data presented here is that the *ospD* gene has only recently been acquired by the Lyme disease spirochetes and that the gene has been laterally transferred among species. This finding argues against the clonal structure of LDS populations as recently proposed by Dykhuisen et al. (13).

MATERIALS AND METHODS

Bacterial isolates and cultivation. Isolates were cultivated in BSKII medium (3) at 34°C, pelleted by centrifugation (7,500 rpm for 15 min in a Beckman JA14 rotor), and washed twice with phosphate-buffered saline (pH 7.0). All of the isolates used are described in Table 1.

DNA isolation and species-specific PCR identification of *Borrelia* isolates. All of the isolates used in this study which had not been previously identified to the species level were classified through the use of LDS species-specific primer sets, which have been previously described (27). The primer sequences are presented in Table 2.

Pulsed-field gel electrophoresis. Sample preparation and transverse alternating-field gel electrophoresis were performed as previously described (30). The DNA was fractionated in 1.0% agarose gels in 0.25× TBE (1× TBE is 89 mM Tris–89

* Corresponding author. Present address: Department of Microbiology and Immunology, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298-0678.

TABLE 1. Description of bacterial isolates and distribution of the *ospD* gene

Species and isolate(s)	Source[a]	Geographic origin	Passage no.	Presence of *ospD* gene
B. burgdorferi				
LP3, LP4, LP5, LP7	Human EM	Connecticut	3, 3, 3, 3	−, −, +, +
NY13, NY1-86	Human EM	New York	15, 20	−, +
272	Human EM	United States	28	−
297	Human CSF	Connecticut	?	−
HBNC	Human blood	California	2	−
3028	Human pus	Texas	H	+
IP2A	Human CSF	France	H	−
CA2-87, CA3-87	*Ixodes pacificus*	California	11, ?	−, −
CA4, CA7, CA8, CA9	*I. pacificus*	California	5, 7, 8, 3	−, −, −, −
DN127, CA12	*I. pacificus*	California	50, 25	−, +
CA13	*I. neotomae*	California	9	−
27985	*I. scapularis*	Connecticut	5	+
B31, Sh2-82, 25015	*I. scapularis*	New York	H, H, 6	+, −, −
21721	*I. scapularis*	Wisconsin	6	−
20004	*I. ricinus*	France	H	−
VS219, VS134	*I. ricinus*	Switzerland	10, 9	−, −
1352	*Amblyomma americanum*	Texas	H	−
Veery	Bird	Connecticut	H	−
Illinois 1	Mouse	Illinois	?	−
JD1	*I. scapularis*	Massachusetts	4	−
T2	*I. scapularis*	?, United States	8	+
B. garinii				
20047, 153	*I. ricinus*	France	?, ?	+, +
VS102, VS492	*I. ricinus*	Switzerland	9, 10	+, −
VS307, VS290	*I. ricinus*	Switzerland	?, 7	+, +
VS/BP	Human CSF	Switzerland	9	+
G25	*I. ricinus*	Sweden	H	+
N34, FRG, PBi	*I. ricinus*	Germany	?, 19, ?	−, +, +
G1, G2	Human CSF	Germany	H, H	+, +
R-IP90, IP89	*I. persulcatus*	Russia	H, 3	+, +
B4-87, B6-91	*I. ricinus*	Norway	?, ?	+, +
B4-91, B5-92	Human EM	Norway	?, ?	+, +
B1-91	Human ACA	Norway	?	+
B. afzelii				
UO1, ECM1, UMO1	Human EM	Sweden	H, H, 15	+, +, −
PGau	Human ACA	Germany	H	−
BO23, PKo	Human EM	Germany	H, H	−, +
VS461	*I. ricinus*	Switzerland	8	+
B1-87	*I. ricinus*	Norway	?	+
R-IP3, R-IP21	*I. persulcatus*	Russia	H, H	−, −
J1	*I. persulcatus*	Japan	?	−
PBo	Human CSF	Germany	12	+
B. japonica sp. nov.				
HO14	*I. persulcatus*	Japan	?	+
IKA2	*I. ovatus*	Japan	?	−
Genospecies 21038				
19857	Rabbit kidney	New York	7	−
21038	*I. dentatus*	New York	6	−

[a] EM, erythema migrans; CSF, cerebrospinal fluid; ACA, acrodermatitis chronica atrophicans.

mM borate–2 mM EDTA) at 13°C by using the GeneLine II system (Beckman). Low-range pulsed-field gel markers (New England Biolabs) and *Saccharomyces cerevisiae* YNN 295 chromosomal DNA (Beckman) were used as molecular weight standards. After staining with ethidium bromide, the DNA was transferred onto Hybond-N membranes by vacuum blotting and probed with an *ospD* amplification product obtained by amplification of the *ospD* gene from *B. burgdorferi* B31 with the F3-R3 primer set as described below. Henceforth, we will refer to this probe as the F3R3 probe. In all cases in which this probe was used, it was labeled by random-primer labeling with the Random Prime labeling kit (Boehringer Mannheim) and

TABLE 2. Oligonucleotide probes and primers

Probe[a]	Sequence (5′ to 3′)	Target site, specificity[b] (positions)
F1(+)	TAGCATCATTAAACATCC	*B. burgdorferi ospD*
F3(+)	ATTGCTCTCAATATCTTG	*ospD*, all LDS species (42–59)
SK2(+)	CTTTAAGTATTAACGGCTG	*B. garinii-B. afzelii ospD*
R1(−)	TGTGCTGGCTTATCCTTC	*B. burgdorferi ospD* (843–860)
R3(−)	GTGGCCTTGTTAAATACTTAA	*ospD* gene, all LDS species (781–801)
R6(−)	CTTGTGTTCATGATAAACAAG	*B. burgdorferi ospD* (56–76)
R8(−)	GAAGCAGTTATAAGTGCGG	*B. garinii-B. afzelii ospD* (277–295)
BB-F(+)	GGGATGTAGCAATACATTC	*B. burgdorferi* 16S rRNA gene (74–92)
BB-R(−)	ATATAGTTTCCAACATAGG	*B. burgdorferi* 16S rRNA gene (630–648)
BG-F(+)	GGGATGTAGCAATACATCT	*B. garinii* 16S rRNA gene (74–92)
BG-R(−)	ATATAGTTTCCAACATAGT	*B. garinii* 16S rRNA gene (630–648)
BA-F(+)	GCATGCAAGTCAAACGGA	*B. afzelii* 16S rRNA gene (59–76)
BA-R(−)	ATATAGTTTCCAACATAGC	*B. afzelii* 16S rRNA gene (630–648)

[a] All primers with the suffix (+) are plus-strand primers, and those with the suffix (−) are minus-strand primers (i.e., reverse complements of the gene sequence). The primers which target the *B. afzelii* 16S RNA gene were previously referred to as VS primers (27) but have been renamed for consistency with recent revisions in the nomenclature of the LDS species (11).

[b] The target sites for the F1(+) and SK2(+) primers occur in the 5′ noncoding region of the *ospD* gene, upstream of the repeat elements. Because of the variability in the number of repeats in that region, the target sites for these primers were not numbered. For other *ospD*-targeting primers, all numbering is in reference to that presented in the alignment in Fig. 3. For the 16S rRNA gene primers, all numbering is in reference to *Escherichia coli*.

[α-^{32}P]CTP (3,000 Ci/mmol). The sequences of the F3 and R3 primers and all of the other primers and probes used in this study are listed in Table 2.

Cloning of the *ospD* gene. DNA from *B. garinii* isolate R-IP90 was restricted with *Hin*dIII, and a portion of the material was fractionated in a 0.8% agarose gel. The DNA was transferred by vacuum blotting onto a Hybond-N membrane (Amersham) and probed with the F3R3 probe. The size of the *ospD*-carrying fragment was determined by Southern blotting, and the appropriate-size fragment was eluted from a second low-melting-point agarose gel with the Magic mini-prep kit (Promega). The isolated fragment was then ligated into the *Hin*dIII site of the pBluescript II-SK(+) vector (Stratagene) by standard methods. Recombinants were screened for the *ospD* gene by colony lifts with the F3R3 probe.

PCR amplification and sequence analysis of the *ospD* gene. The F3R3 probe used in the Southern blots was generated by PCR amplification of the *ospD* gene from 30 ng of *B. burgdorferi* B31 DNA with the F3-R3 primer set (50 pmol). After 25 rounds of amplification under cycle conditions of 95°C for 30 s, 50°C for 20 s, and 72°C for 45 s, 1 μl of the reaction (1%) was withdrawn and used as the template in a second reaction with the same conditions. The reactions and buffers used were as described above. This was done to ensure that the amount of contaminating template DNA would be minimal in the final probe preparation. The amplification product was precipitated from the PCR by incubation for 10 min at room temperature in 1 M NH_4-acetate and 50% isopropanol, followed by centrifugation in a microcentrifuge at full speed. The pellet was dried and resuspended in H_2O to the desired concentration.

ospD sequencing templates from *B. burgdorferi* isolates were generated with the F1-R1 primer set. This primer pair amplifies the entire coding sequence and approximately 250 nucleotides upstream of the translational start codon. The sequence of the *ospD* gene from *B. garinii* isolate R-IP90 was determined by sequencing of the insert with primers targeting the flanking region of the multiple cloning site. On the basis of the sequence determined, additional PCR primers were made which were used to amplify the *ospD* gene in both *B. garinii* and *B. afzelii* isolates. The *ospD* gene amplification products were sequenced with the double-stranded DNA cycle sequencing system (Gibco-BRL).

Sequences were aligned by using a variety of programs, including the PCGENE package and the ALIGN-PLUS package. Some manual refinement of the alignments was performed. Dendrograms were constructed by using the CLUSTAL program contained in the PCGENE package. Gene trees were constructed by using masked sequences and the PHYLIP software package (15). Percent sequence similarity values were determined by using the DNADIST program, and the distances calculated were used to construct a gene tree by using the FITCH program. The confidence interval for each branch node was assessed by bootstrap analysis with the DNABOOT program.

Nucleotide sequence accession numbers. The *ospD* sequences have been assigned GenBank accession numbers U05326 (*B. burgdorferi* LP5), U05327 (*B. burgdorferi* LP7), U05324 (*B. burgdorferi* CA12), U05304 (*B. burgdorferi* 3028), U05305 (*B. burgdorferi* 27985), U05306 (*B. garinii* PBi), U05337 (*B. garinii* G1), U05325 (*B. garinii* R-IP90), U05323 (*B. garinii* 20047), U05328 (*B. afzelii* UO1), U05329 (*B. afzelii* VS461), and U05322 (*B. afzelii* PGau).

RESULTS

Determination of the species identities of LDS isolates. All isolates not previously identified to the species level were classified by use of species-specific rRNA-encoding gene-directed PCR primer sets (27) and/or by restriction fragment length polymorphism (RFLP) pattern analysis of the rRNA-encoding genes as previously described (25) (data not shown). The primer sequences are presented in Table 2. In all cases, the 16S rRNA-encoding gene of each isolate was amplified by only one of the species-specific primer sets. The species assignments inferred from rRNA-encoding gene RFLP pattern analyses were consistent with those obtained via PCR analysis and are in agreement with those reported by others (2). The species identities of all of the isolates tested are listed in Table 1.

Distribution of the *ospD* gene among Lyme disease isolates. All isolates were initially screened for the presence of the *ospD* gene by PCR amplification with the F3-R3 primer set. The sequences of these primers, as well as those of all additional primers and oligonucleotide probes used in this study, are listed in Table 2. Preliminary investigations established that this primer set was suitable for amplifying the *ospD* gene in

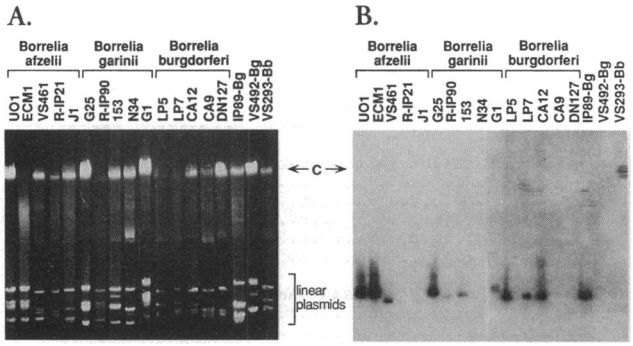

FIG. 1. Analysis of the distribution of the *ospD* gene among Lyme disease isolates by transverse alternating-field gel electrophoresis. Samples were prepared and fractionated in a 1% gel by transverse alternating-field gel electrophoresis as described in the text, stained with ethidium bromide, and visualized (A). The gel from panel A was blotted onto a membrane and probed with the *ospD*-directed F3R3 probe (B). The identity of each sample is indicated above each lane. C, chromosome.

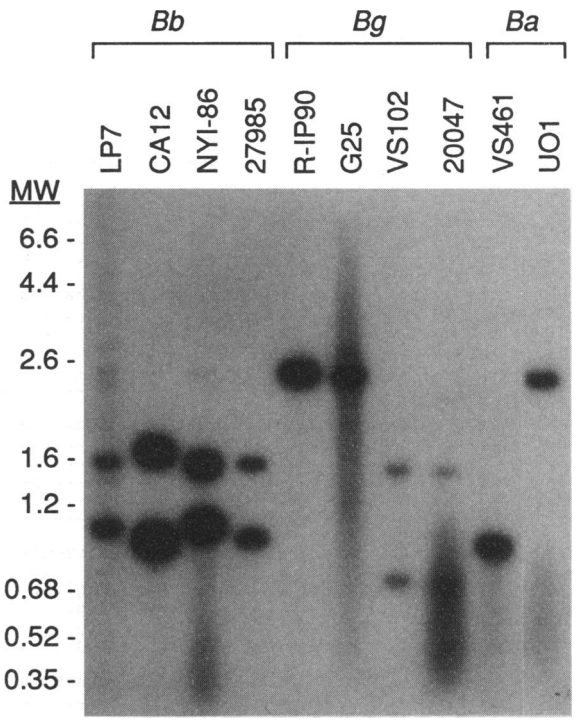

FIG. 2. RFLP patterns for the *ospD* gene. Isolated DNA was digested with *Hin*dIII, fractionated in a 0.8% agarose gel, vacuum blotted onto a Hybond-N membrane (Amersham), and probed with the F3R3 probe as described in the text. The molecular weight (MW) markers indicated on the left (kilobases) are a combination of pGEM (Promega) and lambda *Hin*dIII-digested DNA (Boehringer Mannheim) markers. The identity of each sample is indicated above each lane. *Bb*, *Bg*, and *Ba* are abbreviations for *B. burgdorferi*, *B. garinii*, and *B. afzelii*, respectively.

each of the three LDS species. The results obtained via PCR were verified by Southern blotting with a PCR-generated *ospD* probe obtained with the F3-R3 primer set (Fig. 1). Results concerning the distribution of the gene among isolates are summarized in Table 1. While members of each of the three LDS species possess the gene, not all isolates were found to carry it. The percentage of isolates tested that were found to carry the gene varied depending on the species. We found that 90% of the *B. garinii*, 50% of the *B. afzelii*, and 24% of the *B. burgdorferi* isolates studied carry the gene. The G test of independence showed these differences to be significant ($P < 0.005$). The gene was also detected in one of two *B. japonica* sp. nov. isolates. From a geographic viewpoint, the *ospD* gene was found in 66% of European-Asian isolates and in 25% of North American isolates. The *ospD* gene was not detected in *B. hermsii*, *B. anserina*, *B. coriaceae*, *B. turicatae*, or *B. parkeri*.

The size of the plasmid carrying the *ospD* gene in *B. garinii* and *B. afzelii* isolates was determined by transverse alternating-field gel electrophoresis and subsequent Southern blotting, by probing with various *ospD*-derived PCR probes. Consistent with that originally described by Norris et al. (32) for *B. burgdorferi* isolate B31, the gene was found to reside on a linear plasmid of approximately 38 kb. However, in other isolates, the size of the plasmid carrying the gene ranged from about 36 to 40 kb (Fig. 1).

RFLPs in the *ospD* gene. The RFLP patterns for the *ospD* gene in *Hin*dIII-restricted DNA were found to be variable, even among isolates of a given species (Fig. 2). In some cases, isolates of different species (i.e., *B. garinii* R-IP90 and *B. afzelii* UO1) had identical RFLP patterns. Hence, in contrast to those observed for the 16S and 23S rRNA (2, 25), *fla*, HSP, and OspAB (40) genes, the RFLP patterns of the *ospD* gene do not allow differentiation of the LDS species.

Cloning and sequence analysis of the *ospD* gene. To generate *ospD* sequencing templates, a series of primer sets based upon the *B. burgdorferi* B31 *ospD* gene sequence were designed (32). All primer sets tested were effective for amplifying the *ospD* gene from all *B. burgdorferi* isolates that carry the gene. The F1-R1 set generated a PCR product of approximately 1,071 nucleotides which allowed determination of the entire *ospD* gene sequence from *B. burgdorferi* isolates. However, by using several different primer sets, we were unable to amplify the entire gene and its upstream control elements from *B. garinii* and *B. afzelii* isolates.

To sequence the upstream and downstream regions of the gene from all isolates, we cloned the *ospD* gene from a representative *B. garinii* isolate, R-IP90. Upon screening of potential recombinant *ospD*-carrying plasmids, a single positive was identified. This plasmid was designated pBSD. The plasmid was found to contain an insert of approximately 2,600 bp. The terminal ends of the insert were sequenced by using primers which flanked the polylinker region of the parental plasmid. Homology with the 5′ upstream region of the *B. burgdorferi* B31 *ospD* sequence was found approximately 100 bases into the insert. On the basis of the sequence obtained from the insert, we then generated primers which were found to be effective for amplification of the *ospD* genes from *B. afzelii* and other *B. garinii* isolates. These primers were used to obtain PCR-generated sequencing templates. The *ospD* coding sequences are aligned in Fig. 3. Specific features of the sequences are discussed in more detail below.

Polymorphisms in the *ospD* gene. Initial sequencing results demonstrated that differing numbers of a 17-bp repeat element occur 5′ of the coding sequence. Norris et al. (32) had demonstrated that in the *B. burgdorferi* B31 *ospD* gene, 7 of these 17-bp repeat elements were present. To assess the number of these tandem repeats among isolates, the upstream region of the gene was subjected to PCR amplification with the SK2-R8 primer set for *B. garinii* and *B. afzelii* isolates (Fig. 4) and the F1-R6 primer set for *B. burgdorferi* isolates (data not

```
                1
Bb-B31    gtataaatAT Gaaaaaatta ataaaaatac tactgttaag tttattttta ttgctctcaa tatcttgtgt tcatgataaa caagaattat catcaaaatc
Bb-LP7    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bg-IP90   .......... .......... .........t ......c... .......... c.a....... ........-- ---------- ---------- ----------
Bg-G1     .......... .......... .........t ....c.c... .......... c.a....... ........-- ---------- ---------- ----------
Ba-UO1    ......---- ------.... .........t ......c... .......... a.a....... ........-- ---------- ---------- ----------
Ba-VS461  .......... .......... .........t ......c... .......... c.a....... ........-- ---------- ---------- ----------
              101
Bb-B31    taatttaaat aatcaaaaag gatatttaga taatgaaggc gcaaattcaa a---ttacga atcaaaaaaa cagagcatat taagtgagtt aaatcagtta
Bb-LP7    .......... .......... .......... .......... .......... .---...... .......... .......... .......... ..........
Bg-IP90   ---------- ---------- --.c...... .........t .t........ .aga...... .......... .....t.... ..g....a.. ..........
Bg-G1     ---------- ---------- --.c...... .........t .t........ .aga...... .......... .....t.... ..g....a.. ..........
Ba-UO1    ---------- ---------- --.c...... .........t .t........ .aga...... g......... .....t.... ..g....a.. .......c..
Ba-VS461  ---------- ---------- --.c...... .........t .t........ .aga...... .......... .....t.... ..g....a.. ..........
              201
Bb-B31    ttaaagcaaa ctacaaattc actaaaagaa gccaaaaata caacagataa tttaaatgca tcaaatgagg caaataaagt tgtagaagcg gttataaatg
Bb-LP7    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bg-IP90   ..ggg..... .......... .......c.. ..a....... .......... .......... .......... .......... .........a .......g..
Bg-G1     ..ggg..... .......... .......... ..a.......g....... .......... ....c..... .......... .........a .......g..
Ba-UO1    ..ggg..... .......... .......... ..a....... .......... .......... .......... .......... .........a .......g..
Ba-VS461  ..ggg..... .......... .......... ..a....... .......... .......... .........c a......... .........a .......g..
              301
Bb-B31    cagttaattt aatttcatct gctgcagatc aagtaaaaag tgcaacaaaa aatatgcatg atttagctca aatggcagaa atagatttag aaaaaataaa
Bb-BLP7   .......... .......... .......... .......... ....ca.... .......... .......... .......... .......... ..........
Bg-IP90   .g........ .......... .......... .g.......c. ........c. .......... .......... .......... .......... ..........
Bg-G1     .g........ .......... .......... .g......g. .......... .......... .......... .......... .......... ..........
Ba-UO1    .g........ .......... .......... .g........ ........c. .......... .......... .......... .......... ..........
Ba-VS461  tg........ .......... .......... .g........ ........c. ........c. .......... .......... .......... ..........
              401
Bb-B31    gaactctagt gataaagcaa tatttgcatc taatcttgca aaagaagcat atagccttac taaagcagca gaacaaaaca tgcaaaaact gtataaagag
Bb-LP7    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bg-IP90   .g........ .......... .......gg. ....g..... .......... ...a...... .......... ........t. .......... ..........
Bg-G1     .g.a...... .......t.. .......gg. ....g..... .......... ...a...... .......... ........t. .....g.... ..........
Ba-UO1    .g.a...... .......... ..g....gg. ....g..... .....g.... ...a...... .......... ........t. .......... ..........
Ba-VS461  .g.a...... .......t.. ..g....gg. ....g....g .......... ...a...... ........t. ........t. .......... ..........
              501
Bb-B31    caacaaaaaa tatcagaatc agaatcagaa tctgactatt ctg------- ---------- ---------- attctgctga aataaaacaa gctaaagagg
Bb-LP7    .......... .......... .......... .......... ...------- ---------- ---------- .......... .......... ..........
Bg-IP90   ...g..g.gc a.ct.a..a. .ct...t..t .....tg.aa t--------- ---------- ---------- ---------- ---....... ..........
Bg-G1     ..-------- -.ct.a..a. .ct...t..t .....tg.aa ta.aacaagt ---------- ---------- -.....a... .......... ..........
Ba-UO1    ...g..g.gc a.ct.a..a. .ct...t..t .....tg.aa t--------- ---------- ---------- ---------- ---....... ..........
Ba-VS461  ...g..g.gc a.ct.a..a. .ct...t..t .....tg.aa t.gaacaagt ttctgatgaa ataaaacaag t.....a... .......... ..........
              601
Bb-B31    ccgtagaaat agcttggaaa gctacagtag aagcaaaaga taagttaatt gatgtagaaa atacagtcaa agagacattg gataaaataa agacagaaac
Bb-LP7    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bg-IP90   .t........ .......... .........a .......... .g........ .....g.... ..gt...... ....g..... .......... ..........
Bg-G1     .t........ .......... .....c...a .......... .g........ ........g. ..g....... ....g..... .......... ....t.....
Ba-UO1    .t........ .......... .........a .......... .g........ .....g.... ..gt...... ....g..... .......... ..........
Ba-VS461  .t........ .......... .........a .......... .g........ .......... ..g....... ....g..... .......... ..........
              701
Bb-B31    tacgaacaat acaaagcttg cagatataaa agaagcagca gagttagtat tacaaatagc taagaatgca aaggaaatag tacaagaagt tgtggccttg
Bb-LP7    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bg-IP90   .g........ .....a.... ........g. .....t.... .......... .......... c.......t. gc........ cg........ ....a..c..
Bg-G1     .g..... .....a.... ........g. .....t.... .......... .......... c.......t. .......... cg........ .....a.c..
Ba-UO1    .g.t...... .....a.... ........g. .....t.... .......... .......... c.......t. gc........ cg........ ....a..c..
Ba-VS461  cg........ .....a...a ........g. .....t.... .......... .......... c..a....t. gc........ cg........ .......c..
              801
Bb-B31    ttaaatactT AAataaataa aagatggaag taaggaagga agaataattt tgtgctggct tatccttcc
Bb-LP7    .......... .......... ..........
Bg-IP90   .......... ........tt ttat.aacta a.tt.tg.a. tt.gct.aag .t.ttaa.aa ..aat...t
Bg-G1     .......... .....t..tt ttat.aacta
Ba-UO1    .......... .....t..tt ttat.aacta a.tt.tg.a. tt.gct.aag .t.ttaa.aa ..aat...t
Ba-VS461  .......... .....t..tt ttat.aacta g.tt.tg.a. tt.g.t.aag ct.ttaa.aa ..aat...t
```

FIG. 3. Alignment of representative *ospD* coding sequences. Sequences were determined as described in the text. Representative sequences were aligned, with nucleotides identical to the *B. burgdorferi* (Bb) B31 sequence indicated by dots and gap positions indicated by dashes. The translational start and stop codons are in uppercase letters. As discussed in the text, the alignment presented for the region between positions 500 to 570 is one of several possibilities. Bg, *B. garinii*; Ba, *B. afzelii.*

shown). The plus strand or forward primer for each of these primer sets hybridizes with the gene upstream of the repeat elements, and the minus strand or reverse primer hybridizes just 3′ of the repeats. On the basis of the B31 *ospD* sequence, amplification products of approximately 520 bp were expected when the SK2-DR8 primer set was used and products of approximately 300 bp when the F1-R6 primer pair was used. However, the sizes of the amplification products varied widely, particularly those obtained with *B. garinii* and *B. afzelii* isolates. For these species, the products ranged from approximately 350 bp with *B. afzelii* isolate UO1 to approximately 570 bp with *B. afzelii* isolate VS461 (Fig. 4). Sequence analysis of the PCR products demonstrated that the difference in size of the amplification products was the result of differing numbers of the 17-bp repeat, which ranged from 1 in *B. afzelii* UO1 to 12 in *B. garinii* R-IP90. The sequences of the upstream region of the gene are aligned in Fig. 5. With many isolates (*B. afzelii* isolates PGau and UO1 and *B. garinii* isolates G2, VS102, and 153, for example), more than one amplification product was obtained, demonstrating variation in the organization of the upstream region of the gene within a population. The ratio of these different size amplification products varied among isolates. Southern blotting and probing with internal *ospD* oligonucleotide probes confirmed that these products were specific and not spurious amplification products (data not shown).

In addition to the various numbers of repeat elements, several other interesting features were observed upon alignment of the sequences (Fig. 3). Relative to the *B. burgdorferi* B31 *ospD* sequence, all *B. garinii* and *B. afzelii* isolates were found to have a 54-base deletion beginning 61 bases into the

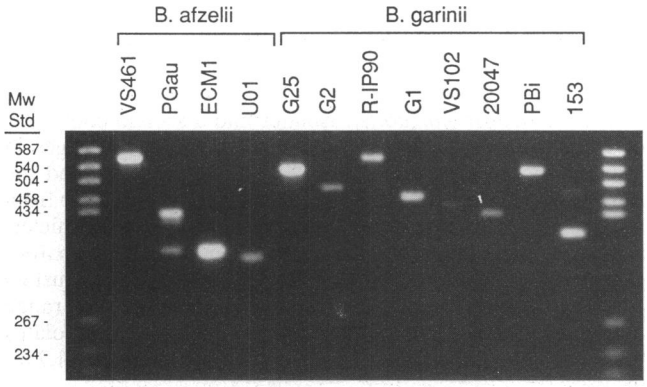

FIG. 4. PCR analysis of the upstream region of the *ospD* gene. The *ospD* genes from representative *B. garinii* and *B. afzelii* isolates were amplified with the SK2-DR8 primer set as described in Materials and Methods, and 15 μl of each reaction mixture was analyzed by electrophoresis in a 2.5% Meta-phor agarose (FMC) gel and ethidium bromide staining. Mw Std, molecular size standards in base pairs.

coding sequence. This deletion does not disturb the open reading frame and results in an OspD protein in these species that is truncated by 18 amino acids. Upstream of the repeat elements, all *B. garinii* and *B. afzelii* isolates also share a 3-bp deletion (Fig. 5), as well as a 4-bp insertion 12 bases upstream from the translational start codon. *B. afzelii* isolate UO1 was found to have a 10-base deletion (relative to the *B. burgdorferi* B31 *ospD* sequence), which results in loss of the translational start codon.

On the basis of polymorphisms in the coding sequence of the gene, we delineated four types of variants. At a region in the coding sequence around base 550, numerous different types of insertions, deletions, and repeats were observed relative to the *B. burgdorferi* B31 *ospD* sequence which we refer to as a type 1 variant (Fig. 6). Alignment of this region is particularly difficult because of the presence of these polymorphisms. Several potential alignments are possible, depending upon whether sequence similarity is maximized or gaps are minimized. The rationale for the alignment in Fig. 6 is that the variation in this region likely resulted from recombinational events. Hence, the alignment presented minimizes disruption of the repeat elements. All of the *B. burgdorferi ospD* sequences that we analyzed, regardless of whether the isolates were of European or North American origin, are of type 1. Type 1 variants contain a direct repeat of six nucleotides which borders the site at which the insertion and deletions occur in the other variant types. Type 2 variants, typified by *B. afzelii* UO1 and *B. garinii* RIP90, have a 15-base deletion with no significant repeats in the region (all comparisons are relative to the B31 *ospD* sequence). Type 3 variants, typified by *B. garinii* G1, have a

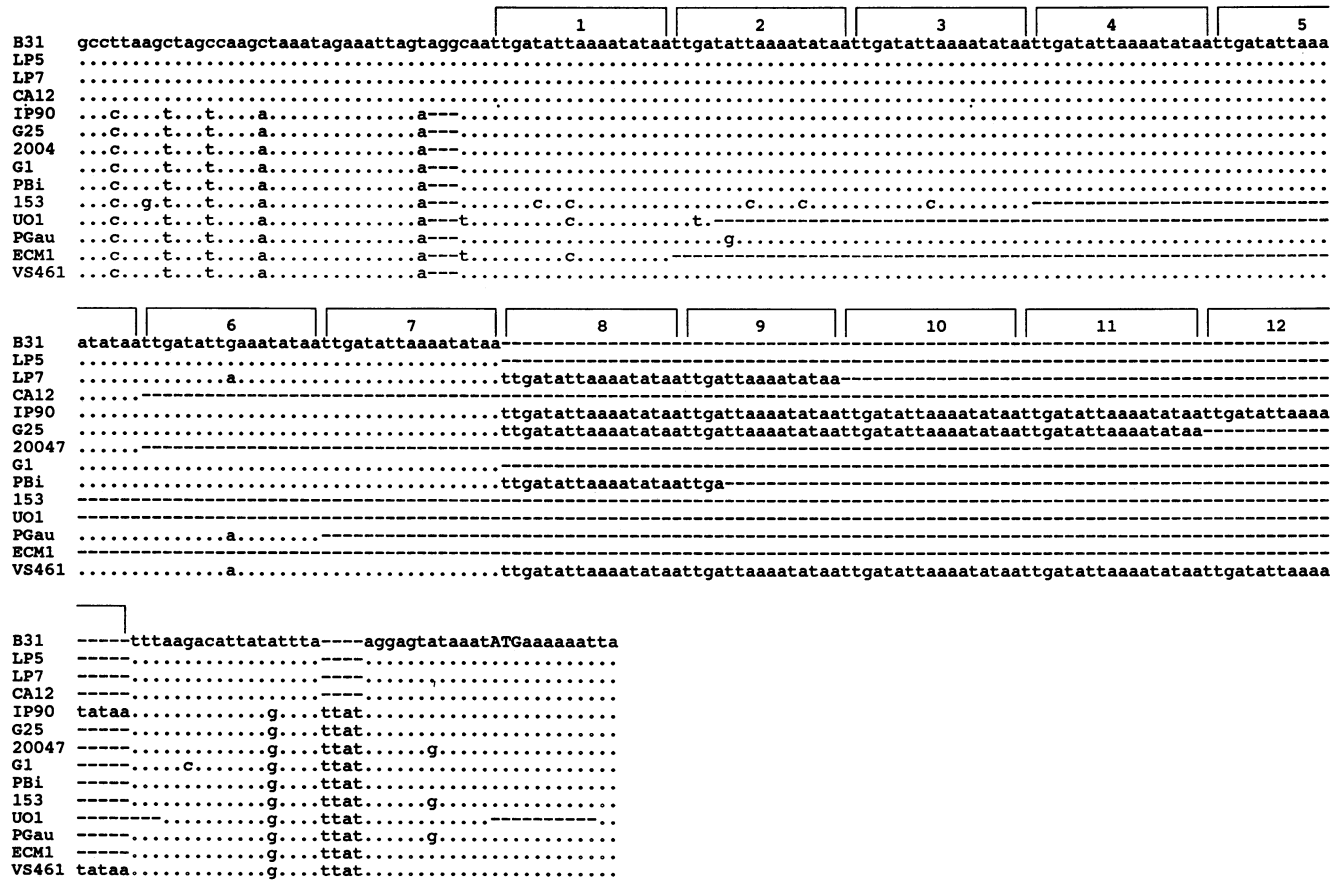

FIG. 5. Alignment of the upstream noncoding regions of *ospD* genes. Repeat elements are indicated by the numbered brackets above the sequence alignment. Positions that are identical to the *B. burgdorferi ospD* B31 sequence are indicated as dots, and gaps are indicated by dashes. B31, LP5, LP7, and CA12 are *B. burgdorferi* isolates; IP90 (referred to as R-IP90 in the text), G25, 20047, G1, PBi, and 153 are *B. garinii* isolates; and UO1, PGau, ECM1, and VS461 are *B. afzelii* isolates.

FIG. 6. Polymorphisms in the *ospD* coding sequence. The *ospD* sequence occurring around base 550 in a representative isolate of each of the four variant types described in the text is presented. Repeat elements are indicated by brackets above the sequences. Gaps are indicated by dashes. The repeated sequences are in uppercase letters, and flanking or mismatch positions are in lowercase letters. The abbreviations Bb, Bg, and Ba stand for *B. burgdorferi*, *B. garinii*, and *B. afzelii* sp. nov., respectively. Several different alignments for this region are possible.

20-base repeat (with one mismatch) separated by a one-nucleotide spacer. Type 4 variants are typified by *B. afzelii* isolates VS461 and B1-87, which have an insertion of 27 bp. In this region, three tandem repeats of 21 nucleotides occur. Other insertions and deletions were observed, and they are depicted in Fig. 3 and 5.

Pairwise sequence comparisons and dendrogram construction. The percent sequence similarities of the *ospD* coding sequences were determined by distance matrix analyses by using masked sequences (i.e., aligned sequences from which gaps and their corresponding positions in sequences lacking gaps are removed). Since the region around nucleotide 550 could not be unambiguously aligned, this region was omitted from the analysis. The coding sequence was highly conserved among *B. burgdorferi* isolates (ranging from 99.4 to 100%), while it was less conserved in *B. garinii* (95.4 to 98.8%) and *B. afzelii* (96.7 to 97.3%). As discussed above, several different alignments for this region are possible. Depending upon the alignment parameters used, the overall similarity values can be increased by an additional 1.5%. This, however, comes at the cost of increasing gap size and number. Presented in Table 3 are the similarity values obtained when pairwise comparisons of the amino acid sequences were performed by using the PALIGN program in the PCGENE package.

The complete deduced amino acid sequences were aligned, and they are presented in Fig. 7. As discussed above, several different alignments for the region around amino acid 185 are possible. These sequences were used to construct a dendrogram (Fig. 8). Two major clusters were identified, one composed solely of *B. burgdorferi* isolates and a second composed of *B. garinii* and *B. afzelii* isolates. While minor changes in branch length did occur, the clustering pattern remained the same when amino acids 170 through 190 were omitted from the analyses. The *ospD* gene tree constructed from the nucleic acid sequences yielded the same result, and bootstrap analysis completely (100%) supported the division into two clusters (data not shown). The OspD proteins from *B. burgdorferi* range in percent identity from 98.8 to 100%, while those from isolates found in the *B. garinii-B. afzelii* cluster range from 91.7 to 98.7%.

DISCUSSION

Norris et al. (32) demonstrated that several proteins are expressed in a low-passage isolate of *B. burgdorferi* B31 that are not expressed after high passage. They suggested that there may be a correlation between the expression of these proteins and the potential infectivity of an isolate, since the high-passage B31 isolate was not infective. They also demonstrated that one of the proteins, whose expression was associated with low passage, was an outer surface-exposed lipoprotein, which they termed OspD. To aid in defining the potential role that OspD may play in infectivity or virulence, we assessed the evolution, distribution, and expression of the gene encoding this protein among the spirochete species associated with Lyme disease. An understanding of these parameters should prove useful in determining if a correlation exists between expression of OspD and infectivity and in assessing the diagnostic and vaccine potential of OspD. In addition, these analyses will also provide a further test of the recently proposed clonality of the *B. burgdorferi* sensu lato complex (13).

The distribution of the *ospD* gene among Lyme disease spirochete isolates differs for each species. The presence or absence of the gene was assessed by two approaches: (i) PCR analysis with primers cross-reactive with the three LDS species and (ii) Southern blot analysis with *ospD*-specific probes. Of the *B. garinii* isolates tested, 90% (18 of 20) were found to carry the gene, in contrast to only 24% of *B. burgdorferi* (8 of 33) and

TABLE 3. Percent sequence similarity between *ospD* coding sequences and deduced amino acid sequences[a]

Isolate	% Similarity to isolate:														
	B31	CA12	LP5	LP7	NY186	27985	3028	IP90	PBi	G25	G1	20047	PGau	VS461	UO1
B. burgdorferi B31		99.6	99.7	99.6	99.9	99.4	99.6	90.4	90.4	92.0	89.6	88.1	89.8	89.4	90.0
B. burgdorferi CA12	99.2		99.8	100	99.7	99.6	99.7	90.2	90.2	91.7	89.4	87.9	89.6	89.3	89.8
B. burgdorferi LP5	99.6	99.6		99.9	99.6	99.7	99.9	90.0	90.0	91.6	89.2	87.7	89.4	89.4	89.6
B. burgdorferi LP7	99.2	100	99.6		99.7	99.6	99.7	90.3	90.3	91.7	89.4	87.9	89.6	89.3	89.7
B. burgdorferi NY186	99.6	99.6	99.2	99.6		99.6	99.4	91.5	91.6	92.0	89.7	88.2	89.9	89.6	90.0
B. burgdorferi 27985	99.6	99.6	100	99.6	99.2		99.6	89.7	89.7	91.2	88.9	87.4	89.1	88.7	89.2
B. burgdorferi 3028	99.2	99.2	99.6	99.2	98.8	99.6		89.9	89.9	91.4	89.1	87.6	89.2	89.9	89.4
B. garinii IP90	83.9	83.9	83.5	85.1	85.5	84.7	84.3		97.8	97.5	97.2	95.8	97.6	97.9	99.1
B. garinii PBi	84.5	84.5	84.1	84.5	84.9	84.1	83.7	94.9		98.5	98.8	95.7	97.5	96.9	97.2
B. garinii G25	83.9	83.9	83.5	83.9	84.3	83.5	83.1	96.2	97.5		97.9	95.4	97.2	96.9	96.9
B. garinii G1	84.1	84.1	83.7	84.1	84.5	83.7	83.3	94.9	98.7	97.9		95.7	98.5	96.6	96.6
B. garinii 20047	80.6	80.6	80.2	80.6	81.0	80.2	79.8	93.2	92.5	92.1	92.5		97.9	94.9	95.7
B. afzelii PGau	81.8	81.8	81.4	81.8	82.2	81.4	81.0	95.7	94.6	94.2	94.6	97.1		96.7	97.3
B. afzelii VS461	79.9	79.9	79.5	79.9	82.3	79.5	81.1	96.6	95.0	94.6	95.8	91.7	94.6		97.3
B. afzelii UO1	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	

[a] The similarity values for the masked *ospD* coding sequences and complete deduced amino acid sequences are presented in the upper right and lower left quadrants, respectively. All calculations were performed as described in the text. ND, not done.

```
1          Sig. Pep. II                    50                                                  100
IP90   MKKLIKILLSSLFLLLSISCS------------------LDNEGVNSKDYESKKQSILGELNQLLGQTTNSLKQAKNTTDNLNASNEANKVVEAVISAVNL
20047  MKKLIKILLSSLFLLLSISCS------------------LDNEGVNSKDYESKKQSILGELNRLLGQTTNSLKEAKNITDDLNASNEANKVVEVVISAVNL
PGau   MKKLIKILLSSLFLLLSISCS------------------LDNEGVNSKDYESKKQSILGELNRLLGQTTNSLKEAKNITDDLNASNEANKVVEVVISAVNL
VS461  MKKLIKILLSSLFLLLSISCS------------------LDNEGVNSKDYESKKQSILGELNQLLGQTTNSLKEAKNTTDNLNASNEQNKVVEAVISVVNL
G25    MKKLIKILLSSLFLLLSISCS------------------LDNEGVNSKDYESKKQSILGELNQLLGQTTNSLKEAKNTADNLNASTEANKVVEAVISAVNL
PBi    MKKLIKILLSSLFLLLSISCS------------------LDNEGVNSKDYESKKQSILGELNQLLGQTTNSLKEAKNTADNLNASSEANKVVEAVISAVNL
B31    MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
LP5    MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
LP7    MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
CA12   MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
3028   MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
27985  MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
NY     MKKLIKILLLSLFLLLSISCVHDKQELSSKSNLNNQKGYLDNEGANSN-YESKKQSILSELNQLLKQTTNSLKEAKNTTDNLNASNEANKVVEAVINAVNL
       ********* **********                 ***** **  ********* *** ** ******* ***  * **** * ***** **  ***

    101                                   150                                                 200
IP90   ISSAADQVKGATTNMHDLAQMAEIDLEKIKESSDKAIVAANVAKEAYNLTKAAEQNMQKLYKEQEEQLKTLSDSD----------EIKQAKEAVEI
20047  ISSAADQVKGATKNMHDLALVAEIDLEKIKESSDKAIVAANVAKEAYNLLKAAEQNMQKLYKEQEEQLKTLSDSDEIEQA-------SKEIKQAKEAIEI
PGau   ISSAADQVKGATTNMHDLALVAEIDLEKIKESSDKAIVAANVAKEAYNLTKAAEQNMQKLYKEQEEQLKTLSDSDEIEQA-------SDEIKQAKEAVEI
VS461  ISSAADQVKGATTNMHDLAQMAEIDLEKIKESSDKVIVAANVAKEAYNLTKAVEQNMQKLYKEQEEQLKTLSDSDEIEQVSDEIKQVSDEIKQAKEAVEI
G25    ISSAADQVKGATKNMHDLAQMAEIDLEKIKDSSDKAIVAANVAKEAYNLTKAAEQNMQELYKEQEEQLKTLSDSDYLEQ-------VSDEIKQAKEAVEI
PBi    ISSAADQVKGATKNMHDLAQMAEIDLEKIKESSDKAILAANVAKEAYNLTKAAEQNMQELYKEQ---LKTLSDSDEIEQ-------VSDEIKQAKEAVEI
B31    ISSAADQVKSATKNMHDLAQMAEIDLEKIKNSSDKAIFASNLAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
LP5    ISSAADQVKSAQKNMHDLAQMAEIDLEKIKNSSDKAIFASNLAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
LP7    ISSAADQVKSAQKNMHDLAQMAEIDLEKIKNSSDKAIFASNVAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
CA12   ISSAADQVKSAQKNMHDLAQMAEIDLEKIKNSSDKAIFASNVAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
3028   ISSAADQVKSAQKNMHDLAQMAEIDLEKIKNSSDKAIFASNLAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
27985  ISSAADQVKSAQKNMHDLAQMAEIDLEKIKNSSDKAIFASNLAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
NY     ISSAADQVKSATKNMHDLAQMAEIDLEKIKNSSDKAIFASNVAKEAYSLTKAAEQNMQKLYKEQQKISESESESDYSD---------SAEIKQAKEAVEI
       ********* *  ******  ********* **** * * * ***** **** ***** *****       * **            * ***********

    201                                   250
IP90   AWKATVKAKDELIDVENVVKEALDKIKTETANNTKLADIEEVAELVLQIAKNVAEIAQEVVTLLNT
20047  AWKATVKVKDELIDVENAVKVALDKIKTETANNTKLADIEEIAELVLQIAKNVKEIAQEVVALLNT
PGau   AWKATVKAKDELIDVENAVKEALDKIKTETANNTKLADIEEVAELVLQIAKNVKEIAQEVVALLNT
VS461  AWKATVKAKDELIDVENAVKEALDKIKTETANNTKLTDIEEVAELVLQIAKNVAEIAQEVVALLNT
G25    AWKATVKAKDELIDVENAVKEALDKIKTETANNTKLADIEEVAELVLQIAKNVKEIAQEVVALLNT
PBi    AWKATVKAKDELIDVENAVKEALDKIKTETANNTKLADIEEVAELVLQIAKNVKEIAQEVVDLLNT
B31    AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT
LP5    AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT
LP7    AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT
CA12   AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT
3028   AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLEIAKNAKEIVQEVVALLNT
27985  AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT
NY     AWKATVEAKDKLIDVENTVKETLDKIKTETTNNTKLADIKEAAELVLQIAKNAKEIVQEVVALLNT
       ****** *** ****** *** ******** ***** ** * ***** ****  ** **** ****
```

FIG. 7. Alignment of representative deduced OspD amino acid sequences. The nucleotide sequences determined were translated by using the PCGENE program and aligned manually. Gaps are indicated by dashes, and residues identical in all of the sequences are underscored with asterisks.

50% of *B. afzelii* (6 of 12) isolates. The gene was not detected in two *Borrelia* isolates (19857 and 21038) previously considered to be LDS isolates which have recently been found to exhibit unique chromosomal maps (11a) and which make up a phyletic group closely related to, but distinct from, the LDS species (24a). The gene was detected in one of two *B. japonica* sp. nov. (23) isolates. An *ospD* equivalent was not detected in other *Borrelia* species, including *B. parkeri*, *B. turicatae*, *B. hermsii*, *B. anserina*, and *B. coriaceae*. Hence, unlike *ospC*, which appears to be widely distributed among several species of the genus *Borrelia* (30), *ospD* is confined to the LDS species. A correlation between the number of times an isolate had been passed in vitro or between the origin of an isolate, i.e., human versus tick, and the presence or absence of the *ospD* gene was not observed.

The distribution of the *ospD* gene differs markedly from that observed for OspA, OspB, and OspC. The *ospAB* operon, which is carried on linear plasmids ranging in size from 50 to 56 kb (5, 36), has been found to be universal among LDS isolates. Some exceptions, however, have been reported, such as in *B. garinii* isolate VS102, which has a deleted *ospB* gene (28), and in several cloned *B. burgdorferi* variants with various recombined *ospAB* operons (34). Other examples include isolates which, under in vitro cultivation and selection, have lost the

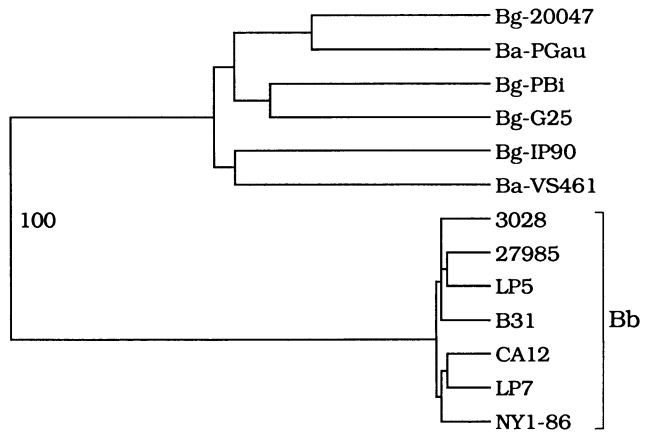

FIG. 8. Dendrogram of translated *ospD* sequences. The dendrogram was constructed by using the CLUSTAL program contained in the PCGENE software package. The number at the node indicates the results of bootstrap analyses performed as described in the text. The abbreviations Bb, Bg, and Ba stand for *B. burgdorferi*, *B. garinii*, and *B. afzelii*, respectively.

plasmid which carries the genes for OspA and OspB (35). In contrast to other *osp* genes, the *ospC* gene is carried on a 26-kb circular plasmid (29, 35). While expression of the *ospC* gene is variable, the gene has been found to be universal among LDS isolates (29). It is difficult to conclude on the basis of the collection of isolates analyzed in this study whether the distribution of the *ospD* gene is influenced by geography or by selective pressures unique to a given species.

Norris et al. (32) demonstrated that in *B. burgdorferi* B31 the *ospD* gene is carried on a 38-kb linear plasmid. Here we demonstrate that the size of the *ospD*-carrying plasmid ranges from approximately 36 to 40 kb. The size of the plasmid does not appear to correlate with the species identity of an isolate. It is unclear if lack of the *ospD* gene in a given isolate strictly correlates with the absence of the plasmid which typically carries the gene. For example, in *B. afzelii* isolate VS461, the gene is carried on an approximately 36-kb plasmid. A plasmid of equivalent size is carried by *B. afzelii* isolate J1, which lacks the *ospD* gene, but it is unclear if these are actually similar plasmids.

An interesting feature of the *ospD* gene identified by Norris et al. (32) was the presence of seven contiguous 17-bp direct repeats in the upstream control region of the gene. These repeats contain a potential −35 promoter sequence. When this region of the gene was PCR amplified, the products were found to range widely in size, suggesting that the number of tandem repeats among isolates may vary. This was confirmed by sequence analysis. The number of repeats ranged from 1 to 12. In most cases, the repeats are direct repeats but some point mutations were observed. One might expect that repeat elements of this sort arose from recombinational events, and we addressed this question by sequencing the entire gene from 20 isolates and determined partial sequences from another 11 isolates. The coding sequences were used to construct gene trees to assess the evolution of the gene and to determine if the number of repeats reflected the patterns of divergence of the coding sequence. No correlation between the repeat number and the clustering pattern in the gene trees was observed, suggesting that the repeat number reflects a recent recombinational event rather than a long-term evolutionary one. While we favor this hypothesis, we cannot exclude the possibility that these repeats arose from some sort of slip strand process during replication of the plasmid. Recently, several examples of recombinational events in the *ospAB* operon have been described (34). Recombinant forms of the *ospAB* operon were detected by screening of individual colonies. Since these recombined forms have not been identified as stable populations in nature, it is unclear if organisms with recombined *ospAB* operons can be maintained and transmitted in a natural setting or whether they arise only upon in vitro cultivation and selection. *ospD* recombination differs in this respect in that the recombined forms represent major proportions of the culture population in many low-passage isolates. This is not surprising in view of the distribution results outlined above, which suggest that *ospD* is not essential for survival either in vivo or in vitro and hence there would be little apparent selective pressure to maintain the ancestral form of the gene.

A second region subjected to apparent recombination occurs within the coding sequence in the area of nucleotide position 550 (numbering is in reference to that in Fig. 3). This region is characterized by several different types of insertions, deletions, and repeats relative to the B31 sequence. The alignment presented for this region represents just one of several possible alignments. The presence of repeat elements bordering this variable region suggests that these polymorphisms resulted from homologous recombination. While the sequence in the region of these insertions and deletions is variable among isolates, in all cases the reading frame is maintained. In view of the high degree of conservation of this gene among the LDS species, such a high degree of recombination within the coding sequence seems paradoxical.

Other nonconserved deletions are found in the *ospD* sequences of various isolates. None of these disrupts the normal reading frame. One deletion of interest which occurs in the *B. afzelii* UO1 isolate results in loss of the normal translational start codon that is used in other isolates. The *ospD* gene in this isolate also appears to lack a functional promoter. While a possible −35 sequence is present, a −10 element is not. A similar deletion occurs in *B. afzelii* ECM1. No *ospD* transcript was detected in either isolate by either primer extension and Northern (RNA) blot analyses, providing indirect evidence for the lack of a functional promoter.

All gene trees constructed from 16S rRNA (25–27), *ospA* (13, 40), *ospC* (39, 43), p93, and flagellin gene sequences (13) indicate that Lyme disease-causing spirochetes are divided into three distinct phyletic groups. These findings are consistent with phylogenies inferred from arbitrarily primed PCR (41), DNA-DNA hybridization (2), 16S rRNA signature nucleotide analysis (1, 26, 27), fatty acid profiles (24), and multilocus enzyme electrophoresis (8). However, the gene tree constructed from the *ospD* sequences resolved only two clusters, which is inconsistent with the studies outlined above. All *B. burgdorferi* isolates were tightly clustered and clearly resolved from *B. garinii* and *B. afzelii* isolates. However, isolates of the latter species clustered together. This is consistent with the relationships inferred from RFLP analysis of the *ospD* gene, which demonstrated that some isolates of different species have the same pattern.

The results presented here provide evidence of lateral transfer of genetic material between LDS species. It seems likely that transfer has occurred readily among strains of *B. garinii* and *B. afzelii*. Since *Hin*dIII sites outside of the gene have remained conserved (as assessed from the RFLP patterns), it is likely that the entire *ospD*-carrying plasmid has been transferred. Further sequence analysis of other regions of the *ospD*-encoding plasmid would prove helpful in assessing this possibility. The tight clustering of *B. burgdorferi ospD* gene sequences and the low distribution of the gene among *B. burgdorferi* strains suggest that the gene was only recently acquired by North American LDS isolates. Recently, Dykhuizen et al. (13) have argued that lateral transfer of DNA between LDS isolates is rare. This conclusion was based upon analysis of the *ospA* (located on a linear plasmid), p93 and *fla* (located on the linear chromosome) genes. However, it is possible that certain genetic elements, such as the *ospD*-carrying plasmid, are transferred more readily than others. It is unclear why transfer of the *ospD* gene or *ospD*-encoding plasmid would occur more readily than that of the *ospAB*-encoding plasmid. It has been widely noted that LDS isolates seldom have identical plasmid profiles, even in a given isolate which has been extensively passed in vitro (37). It is possible that recombination and plasmid transfer play a role in generating highly variable plasmid profiles. One alternative explanation for the high level of *ospD* coding sequence conservation, the variable distribution among isolates, and the apparently high degree of recombination is that the *ospD* gene is carried on or has been introduced into LDS isolates by a phage. Phages have been observed in some LDS isolates by electron microscopy (20, 31). The genome of at least one *B. burgdorferi* phage appears to be an approximately 32-kb linear DNA molecule (35a), which is close to the size of the *ospD*-encoding plasmid.

While the potential of any of the characterized Osps in diagnostics and vaccine development remains to be actualized, the results presented here effectively rule out OspD as a potential candidate. The low distribution of the gene suggests that diagnostic assays based upon an immune response to this protein may lead to a significant number of false negatives. With regard to vaccine development, the low distribution and the apparently high level of variability due to recombination that occur within the *ospD* gene may result in lack of protection. Finally, in view of its low distribution among infective strains, it seems likely that OspD does not represent an essential infectivity determinant. In view of the difference in the distribution of this gene in North America versus Europe, it is tempting to speculate that OspD contributes in some way to the differing clinical manifestations of Lyme disease on these two continents.

ACKNOWLEDGMENTS

We thank P. Rosa and M. Konkel for critical review of the manuscript, Dara Newman for statistical analysis, and Gary Hettrick and Bob Evans for preparing the figures.

REFERENCES

1. **Adam, T., G. S. Gassmann, C. Rasiah, and U. B. Göbel.** 1991. Phenotypic and genotypic analysis of *Borrelia burgdorferi* isolates from various sources. Infect. Immun. **59:**2579–2585.
1a. **Anonymous.** 1992. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. Int. J. Syst. Bacteriol. **42:**656–657.
2. **Baranton, G., D. Postic, I. Saint Girons, P. Boerlin, J.-C. Piffaretti, M. Assous, and P. A. D. Grimont.** 1992. Delineation of *Borrelia burgdorferi* sensu stricto, *Borrelia garinii* sp. nov., and group VS461 associated with Lyme borreliosis. Int. J. Syst. Bacteriol. **42:**378–383.
3. **Barbour, A. G.** 1984. Isolation and cultivation of Lyme disease spirochetes. Yale J. Biol. Med. **57:**521–525.
4. **Barbour, A. G.** 1988. Plasmid analysis of *Borrelia burgdorferi*, the Lyme disease agent. J. Clin. Microbiol. **26:**475–478.
5. **Barbour, A. G., and C. F. Garon.** 1987. Linear plasmids of the bacterium *Borrelia burgdorferi* have covalently closed ends. Science **237:**409–411.
6. **Baril, C., C. Richaud, G. Baranton, and I. Saint Girons.** 1989. Linear chromosome of *Borrelia burgdorferi*. Res. Microbiol. **140:** 507–516.
7. **Bergström, S., V. G. Bundoc, and A. G. Barbour.** 1989. Molecular analysis of linear plasmid-encoded major surface proteins, OspA and OspB, of the Lyme disease spirochaete *Borrelia burgdorferi*. Mol. Microbiol. **3:**479–486.
8. **Boerlin, P., O. Peter, A.-G. Bretz, D. Postic, G. Baranton, and J.-C. Piffaretti.** 1992. Population genetic analysis of *Borrelia burgdorferi* isolates by multilocus enzyme electrophoresis. Infect. Immun. **60:**1677–1683.
9. **Brandt, M. E., B. S. Riley, J. D. Radolf, and M. V. Norgard.** 1990. Immunogenic integral membrane proteins of *Borrelia burgdorferi* are lipoproteins. Infect. Immun. **58:**983–991.
10. **Burgdorfer, W., A. G. Barbour, S. F. Hayes, J. L. Benach, E. Grunwaldt, and J. P. Davis.** 1982. Lyme disease—a tick-borne spirochetosis? Science **216:**1317–1319.
11. **Canica, M., F. Nato, L. du Merle, J. C. Mazie, G. Baranton, and D. Postic.** 1993. Monoclonal antibodies for identification of *Borrelia afzelii* sp. nov. associated with late cutaneous manifestations of Lyme borreliosis. Scand. J. Infect. Dis. **25:**441–448.
11a. **Casjens, S.** Personal communication.
12. **Casjens, S., and W. M. Huang.** 1993. Linear chromosomal physical and genetic map of *Borrelia burgdorferi*, the Lyme disease agent. Mol. Microbiol. **8:**967–980.
13. **Dykhuizen, D. E., D. S. Polin, J. J. Dunn, B. Wilske, V. Preac-Mursic, R. J. Dattwyler, and B. J. Luft.** 1993. *Borrelia burgdorferi* is clonal: implications for taxonomy and vaccine development. Proc. Natl. Acad. Sci. USA **90:**10163–10167.
14. **Eiffert, H., A. Ohlenbusch, W. Fehling, H. Lotter, and R. Thomssen.** 1992. Nucleotide sequence of the *ospAB* operon of a *Borrelia burgdorferi* strain expressing OspA but not OspB. Infect. Immun. **60:**1864–1868.
15. **Felsenstein, J.** 1982. Numerical methods for inferring evolutionary trees. Q. Rev. Biol. **57:**379–404.
16. **Ferdows, M. S., and A. G. Barbour.** 1989. Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. Proc. Natl. Acad. Sci. USA **86:**5969–5973.
17. **Fikrig, E., S. W. Barthold, F. S. Kantor, and R. A. Flavell.** 1990. Protection of mice against the Lyme disease agent by immunizing with recombinant OspA. Science **250:**553.
18. **Fikrig, E., S. W. Barthold, F. S. Kantor, and R. A. Flavell.** 1991. Protection of mice from Lyme borreliosis by oral vaccination with *Escherichia coli* expressing OspA. J. Infect. Dis. **164:**1224–1227.
19. **Fuchs, R., S. Jauris, F. Lottspeich, V. Preac-Mursic, B. Wilske, and E. Soutschek.** 1992. Molecular analysis and expression of a *Borrelia burgdorferi* gene encoding a 22 kDa protein (pC) in *Escherichia coli*. Mol. Microbiol. **6:**503–509.
20. **Hayes, S. F., W. Burgdorfer, and A. G. Barbour.** 1983. Bacteriophage in the *Ixodes dammini* spirochete, etiological agent of Lyme disease. J. Bacteriol. **154:**1436–1439.
21. **Johnson, R. C., G. P. Schmid, F. W. Hyde, A. G. Steigerwalt, and D. J. Brenner.** 1984. *Borrelia burgdorferi* sp. nov.: etiologic agent of Lyme disease. Int. J. Syst. Bacteriol. **34:**496–497.
22. **Jonsson, M., L. Noppa, A. G. Barbour, and S. Bergström.** 1992. Heterogeneity of outer membrane proteins in *Borrelia burgdorferi*: comparison of *osp* operons of three isolates of different geographic origins. Infect. Immun. **60:**1845–1853.
23. **Kawabata, H., T. Masuzawa, and Y. Yanagihara.** 1993. Genomic analysis of *Borrelia japonica* sp. nov. isolated from *Ixodes ovatus* in Japan. Microbiol. Immunol. **37:**843–848.
24. **Livesley, M. A., I. P. Thompson, M. J. Bailey, and P. Nuttall.** 1993. Comparison of the fatty acid profiles of *Borrelia*, *Serpulina*, and *Leptospira* species. J. Gen. Microbiol. **139:**889–895.
24a. **Marconi, R. T.** Unpublished data.
25. **Marconi, R. T., and C. F. Garon.** 1992. Phylogenetic analysis of the genus *Borrelia*: a comparison of North American and European isolates of *B. burgdorferi*. J. Bacteriol. **174:**241–244.
26. **Marconi, R. T., and C. F. Garon.** 1992. Identification of a third genomic group of *Borrelia burgdorferi* through signature nucleotide analysis and 16S rRNA sequence determination. J. Gen. Microbiol. **138:**533–536.
27. **Marconi, R. T., and C. F. Garon.** 1992. Development of polymerase chain reaction primer sets for diagnosis of Lyme disease and for species-specific identification of Lyme disease isolates by 16S rRNA signature nucleotide analysis. J. Clin. Microbiol. **30:** 2830–2834. (Erratum, **31:**1026, 1993.)
28. **Marconi, R. T., M. E. Konkel, and C. F. Garon.** 1993. Variability of the *osp* genes and gene products among species of Lyme disease spirochetes. Infect. Immun. **61:**2611–2617.
29. **Marconi, R. T., D. S. Samuels, and C. F. Garon.** 1993. Transcriptional analyses and mapping of the *ospC* gene in Lyme disease spirochetes. J. Bacteriol. **175:**926–932.
30. **Marconi, R. T., D. S. Samuels, T. G. Schwan, and C. F. Garon.** 1993. Identification of a protein in several *Borrelia* species which is related to OspC of the Lyme disease spirochetes. J. Clin. Microbiol. **31:**2577–2583.
31. **Neubert, U., M. Schaller, E. Januschke, W. Stolz, and H. Schmieger.** 1993. Bacteriophages induced by ciprofloxin in a *Borrelia burgdorferi* skin isolate. Zentralbl. Bakteriol. **279:**307–315.
32. **Norris, S. J., C. J. Carter, J. K. Howell, and A. G. Barbour.** 1992. Low-passage-associated proteins of *Borrelia burgdorferi* B31: characterization and molecular cloning of OspD, a surface-exposed, plasmid-encoded lipoprotein. Infect. Immun. **60:**4662–4672.
33. **Preac-Mursic, V., B. Wilske, E. Patsouris, S. Jauris, G. Will, E. Soutschek, S. Reinhardt, G. Lehnert, U. Klockmann, and P. Mehraein.** 1992. Active immunization with pC protein of *Borrelia burgdorferi* protects gerbils against *Borrelia burgdorferi* infection. Infection **20:**342–349.
34. **Rosa, P. A., T. G. Schwan, and D. Hogan.** 1992. Recombination between genes encoding major outer surface proteins A and B of *Borrelia burgdorferi*. Mol. Microbiol. **6:**3031–3040.

35. **Šadžiene, A., B. Wilske, M. S. Ferdows, and A. G. Barbour.** 1993. The cryptic *ospC* gene of *Borrelia burgdorferi* B31 is located on a circular plasmid. Infect. Immun. **61**:2192–2195.
35a. **Samuels, D. S.** Unpublished data.
36. **Samuels, D. S., R. T. Marconi, and C. F. Garon.** 1993. Variation in the size of the *ospA*-containing linear plasmid, but not the linear chromosome, among the three *Borrelia* species associated with Lyme disease. J. Gen. Microbiol. **139**:2445–2449.
37. **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.
38. **Steere, A. C., R. L. Grodzicki, A. N. Kornblatt, J. E. Craft, A. G. Barbour, W. Burgdorfer, G. P. Schmid, E. Johnson, and S. E. Malawista.** 1983. The spirochetal etiology of Lyme disease. N. Engl. J. Med. **308**:733–740.
39. **Theisen, M., B. Frederiksen, A.-M. Lebech, J. Vuust, and K. Hansen.** 1993. Polymorphism in *ospC* gene of *Borrelia burgdorferi* and immunoreactivity of OspC protein: implications for taxonomy and for use of OspC protein as a diagnostic antigen. J. Clin. Microbiol. **31**:2570–2576.
40. **Wallich, R., C. Helmes, U. E. Schaible, Y. Lobet, S. E. Moter, M. D. Kramer, and M. M. Simon.** 1992. Evaluation of genetic divergence among *Borrelia burgdorferi* isolates by use of OspA, *fla*, HSP60, and HSP70 gene probes. Infect. Immun. **60**:4856–4866.
41. **Welsh, J., C. Pretzman, D. Postic, I. Saint Girons, G. Baranton, and M. McClelland.** 1992. Genomic fingerprinting by arbitrarily primed polymerase chain reaction resolves *Borrelia burgdorferi* into three distinct phyletic groups. Int. J. Syst. Bacteriol. **42**:370–377.
42. **Wilske, B., V. Preac-Mursic, U. B. Göbel, B. Graf, S. Jauris, E. Soutschek, E. Schwab, and G. Zumstein.** 1993. An OspA serotyping system for *Borrelia burgdorferi* based on reactivity with monoclonal antibodies and OspA sequence analysis. J. Clin. Microbiol. **31**:340–350.
43. **Wilske, B., V. Preac-Mursic, S. Jauris, A. Hofmann, I. Pradel, E. Soutschek, E. Schwab, G. Will, and G. Wanner.** 1993. Immunological and molecular polymorphisms of OspC, an immunodominant major outer surface protein of *Borrelia burgdorferi*. Infect. Immun. **61**:2182–2191.