Microbiology (2003), **149**, 1113–1125 DOI 10.1099/mic.0.26120-0

Immunological and genetic characterization of *Borrelia burgdorferi* BapA and EppA proteins

Jennifer C. Miller and Brian Stevenson

Department of Microbiology, Immunology and Molecular Genetics, University of Kentucky College of Medicine, MS 415 Chandler Medical Center, Lexington, KY 40536-0298, USA

Correspondence
Brian Stevenson
bstev0@uky.edu

Received 11 November 2002
Revised 14 January 2003
Accepted 23 January 2003

A large majority of examined Lyme disease spirochaete isolates were demonstrated to contain one or both of the paralogous genes *bapA* and *eppA*. Immunological analyses of serum samples collected from infected patients coupled with comparative sequence analyses indicated that *bapA* gene sequences are quite stable but the encoded proteins do not provoke a strong immune response in most individuals. Conversely, EppA proteins are much more antigenic but vary widely in sequence between different bacteria. Considerable evidence of insertion, deletion and other mutations within *eppA* genes was observed. A number of significant recombination events were also found to have occurred in regions flanking *bapA* genes, while the genes themselves rarely exhibited evidence of mutation, suggesting strong selective pressure to maintain BapA sequences within narrow limits. Data from these and other studies suggest important roles for BapA and EppA during the *Borrelia burgdorferi* infectious cycle.

INTRODUCTION

The spirochaete *Borrelia burgdorferi* contains a strikingly large number of plasmids. The type strain, B31, is known to contain 24 different extrachromosomal DNAs (Casjens *et al.*, 1997, 2000; Fraser *et al.*, 1997; Miller *et al.*, 2000a). The majority of identified infection-associated proteins are encoded on these plasmids, and loss of plasmids has also been associated with reduced infectivity (Labandeira-Rey & Skare, 2001; Purser & Norris, 2000; Schwan *et al.*, 1988; Xu *et al.*, 1996). The location of so many putative virulence determinant genes on plasmids is consistent with the hypothesis that genes conferring advantages for local conditions are generally plasmid-borne, as this makes them more likely to be transferred horizontally and to evolve more rapidly than if they were chromosomally located (Eberhard, 1989). Indeed, all known examples of genetic exchange between Lyme disease spirochaetes have involved plasmid DNAs (this work; Dykhuizen & Baranton, 2001).

Among the many borrelial plasmids is the cp32 family, a group of closely related elements that are generally circular plasmids of about 32 kb in size (Stevenson *et al.*, 2001). All examined Lyme disease spirochaetes contain numerous different members of this plasmid family. The cp32 plasmids are largely identical to each other and vary significantly at only a few loci, one of which is the *erp* locus. The encoded Erp proteins are surface-exposed lipoproteins that are produced during mammalian infection and bind to complement regulatory factor H, apparently contributing to the resistance of these bacteria to host complement-mediated killing (Alitalo *et al.*, 2002; Hellwage *et al.*, 2001; Kraiczy *et al.*, 2001a; Kurtenbach *et al.*, 2002a; Stevenson *et al.*, 2002). One of the first *erp* loci to be identified, the *pG* gene of *B. burgdorferi* strain ZS7, contains an unrelated downstream gene named *bapA* (Wallich *et al.*, 1995). The *erp* genes of two other *B. burgdorferi* strains have since been extensively studied, and both strains B31 and 297 contain an *erp* locus that includes a *bapA* gene (Akins *et al.*, 1995; Stevenson *et al.*, 1996). Northern blot analysis indicated that the *bapA* gene of strain B31 is cotranscribed with the adjacent *erp* gene (Stevenson *et al.*, 1998a). This genetic pairing suggests that BapA may also perform an important function for *B. burgdorferi*. While that role remains unknown, previous immunological evidence indicates that BapA functions during mammalian infection (Bauer *et al.*, 2001).

Some *B. burgdorferi* also contain a gene related to *bapA*, named *eppA*; these two genes together constitute *B. burgdorferi* gene paralogue family 95 (Casjens *et al.*, 2000). Strain B31 is known to carry two different *eppA* alleles, *eppA1* and *eppA2*, each on a different 9 kb circular plasmid, cp9-1 and cp9-2, respectively (Champion *et al.*, 1994; Miller *et al.*, 2000a). Sequence analysis of cp9-1 indicated that it probably evolved from a cp32 plasmid through a series of deletion and rearrangement mutations (Casjens *et al.*, 2000). The only other previously described *eppA* gene is that of strain N40, which is likewise located on a 9 kb circular plasmid (Stewart *et al.*, 2001). An earlier study found that infected laboratory rabbits and at least one tested human Lyme disease patient produced antibodies that recognized recombinant strain B31 EppA1 protein

The GenBank accession numbers for the sequences determined in this work are given in the text.

(Champion *et al.*, 1994), indicating that both members of protein paralogue family 95 are expressed by *B. burgdorferi* during mammalian infection.

Prior to the present study, only a very limited number of *B. burgdorferi* strains had been examined for either *bapA* or *eppA*, so it was unknown how widely spread these genes are among Lyme disease bacteria, or how well conserved are their sequences. Likewise, it was unknown how frequently human Lyme disease patients produce antibodies directed against BapA or EppA. Since serological diagnosis of Lyme disease is often complicated by amino acid differences among antigenic proteins, or by lack of antigenicity by proteins synthesized during human infection, such knowledge would be useful for the development of improved serodiagnostic tests. In addition, the juxtaposition of *bapA* genes in operons with genes encoding factor H-binding Erp proteins raises questions as to the origin of *bapA*. To answer these questions, we studied serum reactivity of numerous Lyme disease patients, and examined the genetics of a broad spectrum of spirochaetes isolated from human patients and wild animals around the globe.

METHODS

Bacterial strains and growth conditions. *B. burgdorferi* strain B31 was originally isolated from an *Ixodes scapularis* tick collected in Shelter Island, NY, USA (Burgdorfer *et al.*, 1982). Since its isolation, numerous laboratories have independently maintained different subcultures of this strain. The genome of subculture B31-MI was recently sequenced (Casjens *et al.*, 2000; Fraser *et al.*, 1997). B31-MI contains both *eppA1* (on cp9-1) and *bapA* (on cp32-3) (Casjens *et al.*, 2000; Fraser *et al.*, 1997), but lacks the *eppA2* allele (Miller *et al.*, 2000a). Bacteria in the subculture B31-RML all contain *bapA* but lack *eppA1*, and many also lack *eppA2* (Miller *et al.*, 2000a). Origins and apparent genotypes of all other *B. burgdorferi* strains utilized in this study are listed in Table 1. Many bacteria were isolated from human tissues, others from natural reservoir vertebrates such as *Peromyscus leucopus* mice, and others from vector ticks *Ixodes scapularis*, *I. ricinus*, *I. persulcatus* and *I. pacificus*.

B. burgdorferi were grown in either Barbour–Stoenner–Kelly II (BSK-II) (Barbour, 1984) or BSK-H (Sigma) medium at 35 °C. Clones were obtained from some *B. burgdorferi* isolates by plating the bacteria in BSK semisolid media (Kurtti *et al.*, 1987; Rosa *et al.*, 1996). Well-isolated colonies were then picked out with a sterile Pasteur pipette and grown in liquid media as above.

Southern blot analysis. *B. burgdorferi* were grown to late exponential phase (approx. 10^7 to 10^8 bacteria ml^{-1}). Plasmid DNAs were purified using Qiagen mini-prep kits. Aliquots (300 ng) of total plasmid DNAs, either uncut or digested with *Eco*RI, were separated by agarose gel electrophoresis and transferred to nylon membranes.

DNA probes derived from the strain B31 *bapA* and *eppA1* genes were generated by PCR amplification of B31-MI or B31-RML genomic DNA, utilizing oligonucleotide primer pairs EA-C plus EA-D and BA-3 plus BA-4, respectively (Table 2). An aliquot of each PCR reaction was examined for amplicon purity by agarose gel electrophoresis and staining with ethidium bromide. PCR products were purified through Centricon-100 concentrators (Ambion) (Stevenson *et al.*, 1996), and labelled with [^{32}P]dATP by random priming (ICN). Each labelled probe was individually incubated with the nylon membranes overnight at 45 °C in 6 × SSC (where 20 × SSC is 3 M NaCl, 0·3 M sodium citrate), 0·1 % SDS and 0·5 % nonfat dry milk. Membranes were washed four times for 15 min each at room temperature in 2 × SSC, 0·1 % SDS (low-stringency wash) then exposed to a phosphoscreen or Kodak X-OMAT film overnight. Phosphoscreen blots were analysed utilizing a STORM phosphoimager (Molecular Dynamics). Blots were next washed further at 55 °C in 0·2 × SSC, 0·1 % SDS (high-stringency wash), then exposed and analysed as above. For reuse, blots were then stripped of hybridized probe by successive washes in boiling water. Successful stripping was confirmed by overnight exposure of the membrane to X-ray film.

Sequencing of *bapA* and *eppA* genes. PCR amplification of genes from *B. burgdorferi* isolates was performed using either purified plasmid DNA or crude bacterial lysates as template, and oligonucleotides listed in Table 2. *bapA* gene amplification used primer pairs BA-1 plus BA-2A or BA-3 plus BA-4. Amplification of *eppA* genes utilized oligonucleotide pairs EA-A plus EA-B or EA-C plus EA-D. DNA sequences flanking the *bapA* genes of some bacterial strains were also amplified by PCR. For those studies, DNA 5′ of *bapA* was amplified using oligonucleotides B3 plus BA2A, while DNA 3′ of the gene was amplified with oligonucleotides BA-1 plus CP-0. Oligonucleotide B3 is complementary to a moderately conserved sequence in cp32 plasmid maintenance loci, and CP-0 is complementary to a highly conserved DNA sequence located approximately 2 kb 3′ of every known *erp* locus (Amouriaux *et al.*, 1993; Stevenson *et al.*, 1997, 1998b). DNA 5′ of the *B. burgdorferi* strain J1 *bapA* locus could not be amplified using the B3–BA2A oligonucleotide pair, but was instead amplified using oligonucleotide BA2A paired with E-181, which is complementary to a sequence in the well-conserved *erp* promoter region. Amplicons were analysed by agarose gel electrophoresis with ethidium bromide staining. Some amplicons were cloned into pCR2.1 (Invitrogen), and transformed into *Escherichia coli*. Resultant plasmids were purified and DNA inserts sequenced. Other amplicons were instead purified using Centricon-100 microconcentrators (Ambion) and sequenced without cloning (El-Hage *et al.*, 1999). For the majority of analysed *bapA* and *eppA* genes, two separate PCR amplicons were sequenced. DNA and protein alignments were prepared utilizing CLUSTAL X (Jeanmougin *et al.*, 1998) and phylogenetic trees were constructed utilizing PAUP* version 4.0b10 (Swofford, 2000).

Recombinant proteins and polyclonal antibodies. Polyhistidine-tagged fusion proteins were produced for use in immunoblot analyses. *B. burgdorferi* B31 *bapA*, *eppA1* and *eppA2* genes were PCR amplified and cloned into pUni/V5-His-TOPO (Invitrogen), and polyhistidine-tagged fusions were produced according to the manufacturer's instructions. Recombinant proteins were then purified over nickel affinity columns (Novagen) and analysed for purity by SDS-PAGE with Coomassie brilliant blue staining. Antisera were raised to recombinant EppA1 and BapA proteins in New Zealand White rabbits (El-Hage *et al.*, 2001). Polyclonal antibodies were then affinity purified using the appropriate recombinant protein immobilized on a nylon membrane (Hefty *et al.*, 2001).

Western blot analysis. Aliquots (50 μg) of each polyhistidine-tagged fusion protein were separated by SDS-PAGE and transferred to nitrocellulose membranes (Bio-Rad). Each membrane was blocked overnight at 4 °C in Tris-buffered saline-Tween 20 (TBS-T) (20 mM Tris/HCl, pH 7·5, 150 mM NaCl, 0·05 % Tween 20) containing 5 % (w/v) non-fat dry milk. Patient serum samples were each diluted 1 : 200 in TBS-T and incubated with membranes in a mini-PROTEAN II multiscreen apparatus (Bio-Rad). Membranes were then washed extensively with TBS-T and incubated with protein A–horseradish peroxidase conjugate (Amersham) in TBS-T. Bound antibodies were detected by enhanced chemiluminescence (Amersham).

Table 1. *B. burgdorferi* strains used in this study, and *eppA* and *bapA* Southern blot and PCR results

Strain	Isolated from	Country or US state of isolation	Genovar*	Source	*bapA* Southern	*bapA* PCR	*eppA* Southern	*eppA* PCR
B31	*I. scapularis*	New York	burgdorferi	Burgdorfer *et al.* (1982)	+	ND†	+	ND
N40	*I. scapularis*	New York	burgdorferi	Barthold *et al.* (1988)	ND	−	ND	+
297	Human CSF	Connecticut	burgdorferi	Steere *et al.* (1984)	ND	+	ND	ND
ZS7	*I. ricinus*	Germany	burgdorferi	Zimmer *et al.* (1990)	+‡	ND	+‡	−
90-1654	Human	Germany	unknown	Miller *et al.* (2000a)	+‡	+	+	−
91-1226	Human	Maryland	unknown	Miller *et al.* (2000a)	+‡	+	+‡	+
93-0107	Human	Wisconsin	unknown	Miller *et al.* (2000a)	+‡	+	−	−
93-1425	Human	Wisconsin	unknown	Miller *et al.* (2000a)	+	−	−	−
91-1223	Human	Wisconsin	unknown	Miller *et al.* (2000a)	+‡	+	+‡	+
92-0953	Human	California	unknown	Miller *et al.* (2000a)	+‡	+	+‡	+
93-0117	Human	Wisconsin	unknown	Miller *et al.* (2000a)	+‡	+	−	+
93-1413	Human	Wisconsin	unknown	Miller *et al.* (2000a)	−	−	−	−
95-0024	Human	Wisconsin	unknown	Miller *et al.* (2000a)	+‡	+	+‡	+
114a	Human	New York	unknown	Miller *et al.* (2000a)	−	−	+	+
116a	Human	New York	unknown	Miller *et al.* (2000a)	+‡	+	−	−
119a	Human	New York	unknown	Miller *et al.* (2000a)	+	−	+‡	+
120a	Human	New York	unknown	Miller *et al.* (2000a)	+‡	+	−	−
122a	Human	New York	unknown	Miller *et al.* (2000a)	+‡	+	−	−
124a	Human	New York	unknown	Miller *et al.* (2000a)	+‡	+	−	−
127b	Human	New York	unknown	Miller *et al.* (2000a)	+‡	+	+	+
B356	Human skin	New York	burgdorferi	Wang *et al.* (2001)	+‡	+	−	−
B348	Human skin	New York	burgdorferi	I. Schwartz, New York Med. Coll., Valhalla	+	−	+	+
B331	Human skin	New York	burgdorferi	I. Schwartz	+‡	+	+‡	−
BL203	Human blood	New York	burgdorferi	I. Schwartz	+‡	+	+	−
BL206	Human blood	New York	burgdorferi	Wang *et al.* (2001)	+	+	+	−
BL268	Human blood	New York	burgdorferi	I. Schwartz	+	+	−	−
B023	Human skin	Germany	afzelii	Xu & Johnson (1995)	+	−	+	+
Bits	*I. ricinus*	Italy	afzelii	Cinco *et al.* (1989)	+	−	+	+
PD89	Human blood	China	garinii	Xu & Johnson (1995)	+‡	+	+	+
PBi	Human CSF	Germany	garinii	Masuzawa *et al.* (1991)	+‡	+	+	+
IP89	*I. persulcatus*	Russia	garinii	Xu & Johnson (1995)	+‡	+	−	−
IP90	*I. persulcatus*	Russia	garinii	Kryuchechnikov *et al.* (1988)	ND	−	ND	−
MM1	*P. leucopus*	Minnesota	burgdorferi	Xu & Johnson (1995)	−	+	+‡	+
NCH-1	Human skin	Wisconsin	burgdorferi	Hughes *et al.* (1992)	+‡	+	+‡	−
10293	Bird	Eastern US	burgdorferi	McLean *et al.* (1993)	+‡	+	+	−
20004	*I. ricinus*	France	burgdorferi	Xu & Johnson (1995)	+	−	+‡	−
CT-1	*I. scapularis*	Wisconsin	burgdorferi	Hyde & Johnson (1988)	+‡	−	+‡	+
Ir210	*I. ricinus*	Russia	garinii	Xu & Johnson (1995)	+‡	+	+	+
PotiB1	*I. ricinus*	Portugal	lusitaniae	Postic *et al.* (1994)	−	−	−	−
PotiB2	*I. ricinus*	Portugal	lusitaniae	Postic *et al.* (1994)	+	+	−	−
VS116	*I. ricinus*	Switzerland	valaisiana	Postic *et al.* (1994)	−	−	−	−
M63	*I. ricinus*	Netherlands	valaisiana	Nohlmans *et al.* (1995)	−	−	−	−
DN127 cl-9	*I. pacificus*	California	bissettiae	Bissett & Hill (1987)	ND	+	ND	ND
PGau	Human skin	Germany	afzelii	Xu & Johnson (1995)	+	+	+	+
ECM-1	Human skin	Sweden	afzelii	Xu & Johnson (1995)	+‡	+	+	+
Sh-2-82	*I. scapularis*	New York	burgdorferi	Schwan *et al.* (1988)	+‡	+	+‡	+
CA6	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	+‡	+
CA15	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	+‡	+
CA17	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	−	+
CA19	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	+‡	+
CA20	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	+‡	+
CA21	*I. pacificus*	California	bissettiae	Schwan *et al.* (1993)	+‡	+	+‡	+
CA22	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	+‡	+

Table 1. cont.

Strain	Isolated from	Country or US state of isolation	Genovar*	Source	*bapA* Southern	*bapA* PCR	*eppA* Southern	*eppA* PCR
CA25	*I. pacificus*	California	unknown	Schwan *et al.* (1993)	+‡	+	+‡	+
VS461	*I. ricinus*	Switzerland	afzelii	Baranton *et al.* (1992)	+‡	−	+	+
J1	*I. persulcatus*	Japan	afzelii	Marconi *et al.* (1993)	+‡	+	−	−

*There is some controversy as to whether genetic differences observed among Lyme disease spirochaetes are due to minor variations within the single species *B. burgdorferi*, or if the differences warrant splitting the bacteria into more than 10 proposed novel species. Until this question is resolved, we prefer the conservative approach of assigning bacteria to different genovars of the single species *B. burgdorferi*.

†ND, Not done.

‡Probe hybridization to these DNAs under high-stringency wash conditions.

Human sera. Pre-existing human serum samples were obtained and used in strict accordance with policies and guidelines established by the US Public Health Service and the University of Kentucky Office of Research Integrity. Patient serum samples designated NY1 through NY20 (Table 3), taken from confirmed Lyme disease patients, were provided by Gary Wormser (New York Medical College, Valhalla, NY, USA) and the time interval between onset of symptoms and serum collection was approximately 30 days (Miller *et al.*, 2000b). Patient serum samples designated 90-2631 through 96-1103 (Table 3) were provided by Martin Schriefer (Centers for Disease Control and Prevention, Fort Collins, CO, USA). The time interval between onset of symptoms and collection of these sera varied and has previously been reported (Miller *et al.*, 2000b). Dr Schriefer also provided six control serum samples collected from healthy people residing in parts of the USA where Lyme disease is non-endemic (Table 3) (Miller *et al.*, 2000b). Lyme disease serum samples designated NY62 through NY137 (Table 3) were collected from patients residing on Long Island, NY (Simpson *et al.*, 1990; Stevenson *et al.*, 1998a). The interval between onset of symptoms and sera collection is unknown for this last set of samples (Simpson *et al.*, 1990; Stevenson *et al.*, 1998a).

Nucleotide sequence accession numbers. The *eppA* and *bapA* nucleotide sequences determined for the *B. burgdorferi* strains described herein have been deposited in GenBank and assigned the accession numbers AF468845–AF468864 and AF475909–AF475936, for the *eppA* and *bapA* alleles, respectively. Nucleotide sequences of the plasmid segregation loci physically linked to the *bapA* genes of strains BL206, 93-0107, DN127cl9-1, 90-1654 and CA15 have been assigned accession numbers AY142102–AY142106.

RESULTS AND DISCUSSION

B. burgdorferi bapA genes

Prior to the present study, only three strains of *B. burgdorferi* were known to contain *bapA* genes. The B31 and ZS7 *bapA* genes are identical to each other, while the strain 297 gene differs only slightly (Fig. 1). To gain a statistically significant perspective on this gene, an additional 53 strains were examined for *bapA*. These included bacteria isolated from human Lyme disease patients, wild vertebrates, and ticks throughout the world.

As a first approach, purified plasmid DNAs were examined by Southern blotting using a DNA probe derived from the strain B31 *bapA*. Low-stringency conditions resulted in hybridization of the probe to DNAs of 44 of 52 tested bacteria (Table 1). At higher-stringency wash conditions, hybridization with 34 of those DNAs could be detected. Electrophoretic mobilities of hybridized DNAs were all consistent with the detected *bapA* genes being located on 32 kb circular plasmids (data not shown). Next, PCR was attempted using oligonucleotide primer pairs derived from the known sequences of the *bapA* genes of strains B31, ZS7 and 297. Amplicons were obtained from 39 of the tested strains. In some cases, primarily involving Eurasian isolates, a

Table 2. Oligonucleotides used in this study

Oligonucleotide	Sequence (5′–3′)
EA-A	AGCCCCTTTATCTCTATTATTGTTAATGGTCTTATTATTTGG
EA-B	CTCGCTTTATATAATGCTATCTATACGCCTCATAAGG
EA-C	AAAAGTGCTTTTTATGAGTCAAG
EA-D	TTTCTATTTTTAAAATTAATCTTTAGG
BA-1	TTTCTCGAGAGTGCAAATATAGAAA
BA-2A	TTTGGATCCTCAAATTAAAATCTAATCTTTG
BA-3	TTTAGGCAATAAAGCTAGGAGGGG
BA-4	AGTCTAATCATATCCTCAGACAGG
B-3	TTTATGGGAAAAAATACCCGG
CP-0	GAAAAAGATAACATGCAAGATACG
E-181	TGCTGTTTGGATAGCCTCGAACTC

Table 3. Results of Western blot analyses of Lyme disease patient serum samples for the presence of antibodies which recognize EppA and BapA proteins

Serum sample	BapA*	EppA1*	EppA2*	Serum sample	BapA*	EppA1*	EppA2*
NY1	+	−	+	90-2631 (90-1654)†	−	+	−
NY2	−	−	−	91-1222 (91-1223)†	−	+	+
NY3	−	+	+	91-1458 (91-1226)†	−	+	−
NY4	−	−	−	92-1245‡	−	−	−
NY5	−	+	+	92-1251‡	−	−	−
NY6	−	+	+	92-1316‡	−	−	−
NY7	−	+	+	92-1318‡	−	−	−
NY8	−	+	+	92-1322‡	−	−	−
NY9	−	−	+	92-1328‡	−	−	−
NY10	+	−	+	92-1682 (92-0953)†	−	−	−
NY11	−	+	+	92-1941	−	−	−
NY12	−	−	−	93-0206 (93-0117)†	−	+	+
NY13	−	−	−	93-0208 (93-0107)†	−	−	−
NY14	−	−	−	93-1414 (93-1413)†	−	−	+
NY15	−	+	−	93-1426 (93-1425)†	−	−	+
NY16	−	−	−	94-0880 (95-0024)†	+	−	+
NY17	−	+	−	96-1049 (114a)†,§	−	−	−
NY18	−	−	−	96-1050 (114a)†	−	+	+
NY19	−	+	−	96-1054 (116a)†,§	−	−	−
NY20	−	+	−	96-1055 (116a)†	−	−	−
NY62	−	−	−	96-1060	−	−	−
NY66	−	−	+	96-1062 (119a)†,§	−	−	−
NY67	−	−	−	96-1063 (119a)†	−	−	−
NY68	−	−	−	96-1068 (120a)†,§	−	−	−
NY69	−	−	−	96-1069 (120a)†	−	−	−
NY70	−	−	−	96-1077 (122a)†,§	−	+	−
NY106	+	−	−	96-1078 (122a)†	−	+	−
NY113	−	−	−	96-1087 (124a)†,§	−	+	−
NY115	+	−	−	96-1088 (124a)†	−	+	−
NY116	+	−	−	96-1102 (127b)†,§	−	+	+
NY137	−	−	−	96-1103 (127b)†	−	+	+

*Western blots were scored as yielding a detectable signal (+) or lacking a signal (−).

†*Borrelia burgdorferi* (sample designation is given in parentheses) were cultured from these patients and DNA extracted from each culture was used in PCR and Southern blot experiments.

‡Control serum samples obtained from healthy human volunteers.

§The serum sample and the one immediately below are paired acute and convalescent samples.

PCR product was obtained only when using the internal oligonucleotide primer pair BA-3–BA-4. Including the three strains known previously to contain *bapA*, 86 % of 56 examined *B. burgdorferi* isolates contain a detectable *bapA* gene.

To assess the degree of diversity among BapA proteins, 28 of the *bapA* amplicons were randomly selected for sequencing. Comparison of the predicted proteins indicated a tremendous degree of conservation among BapA proteins, with identities ranging from 92 to 100 % (Fig. 1A, Fig. 2). All analysed *bapA* genes are predicted to yield proteins having molecular masses of approximately 20 kDa. Most BapA proteins are predicted to have slightly basic pI values of approximately 7·8, although some proteins, such as that of strain 297, have acidic pI values of approximately 5·6. The *bapA* genes of strains 91-1226 and BL203, isolated from Lyme disease patients in Maryland and New York, respectively, were identical to the genes of the New York strain B31 and German strain ZS7. Genes from several other geographically dispersed *B. burgdorferi* isolates were also identical. For example, identical *bapA* genes are carried by strains 297 and Sh-2-82, and seven additional strains isolated from patients living in New York and Wisconsin, and from ticks captured in New York and California. With only two exceptions, all the BapA proteins contained an arginine-glycine-aspartate (RGD) motif at amino acids 66–68 (Fig. 1). Virulence-associated proteins of many other pathogenic bacteria contain such RGD motifs, which are critical for integrin binding that facilitates attachment of bacteria to eukaryotic host tissues (Hynes, 1992).

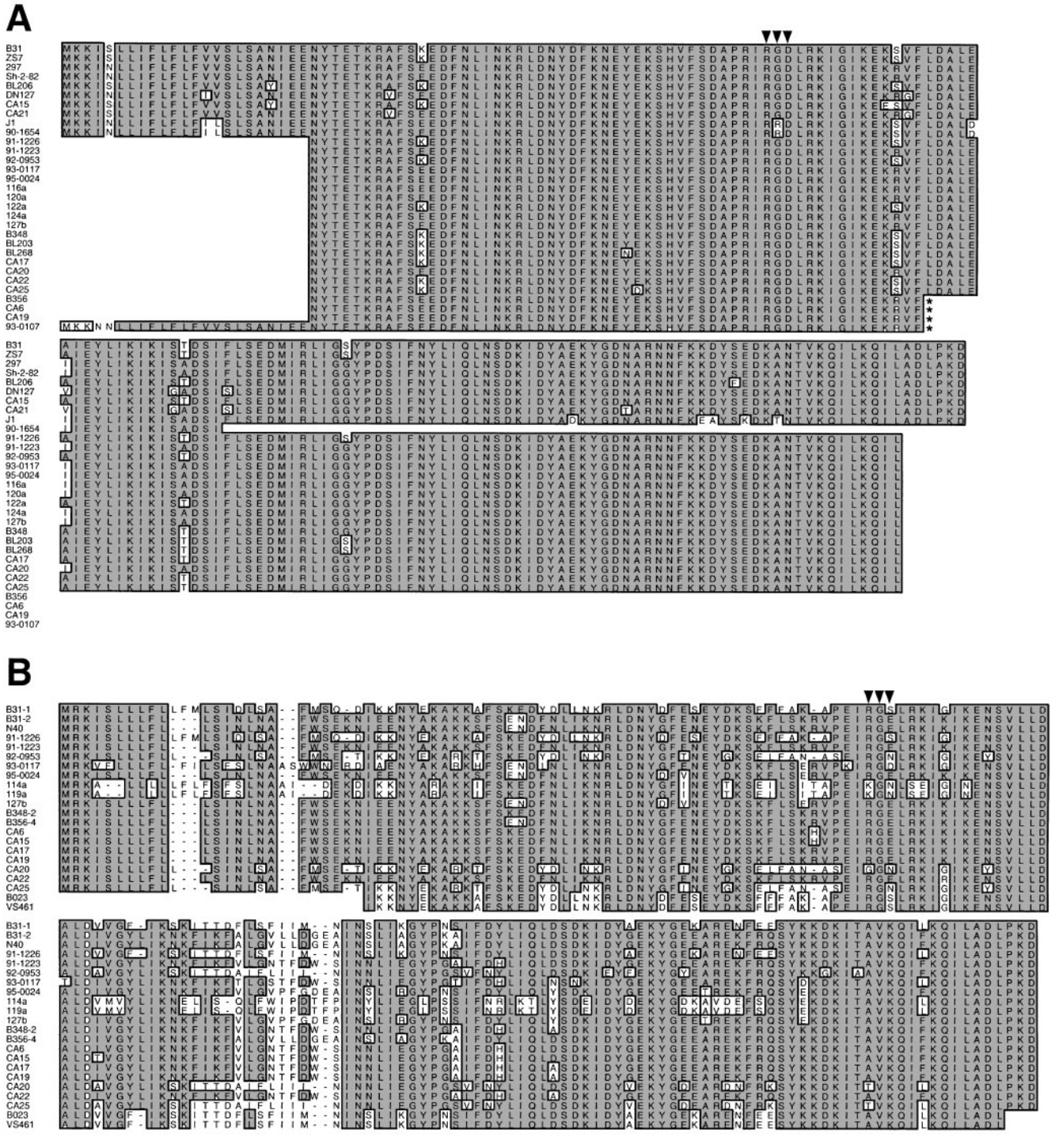

Fig. 1. Alignments of predicted BapA (A) and EppA (B) protein sequences. In panel B, B31-1 refers to the EppA1 protein, and B31-2 to the EppA2 protein. Conserved amino acids are boxed and shaded. RGD motifs of BapA proteins and the paralogous residues of EppA proteins are indicated by arrowheads above each alignment. Gaps introduced to maximize alignments are indicated by dashes. Unknown sequences are left blank. The 93-0107, B356, CA6 and CA19 *bapA* genes all contain a frameshift mutation, and their protein sequences are shown up to the point of truncation, indicated by asterisks. Due to the oligonucleotide primers used for amplifications, only partial sequence was obtained for many genes.

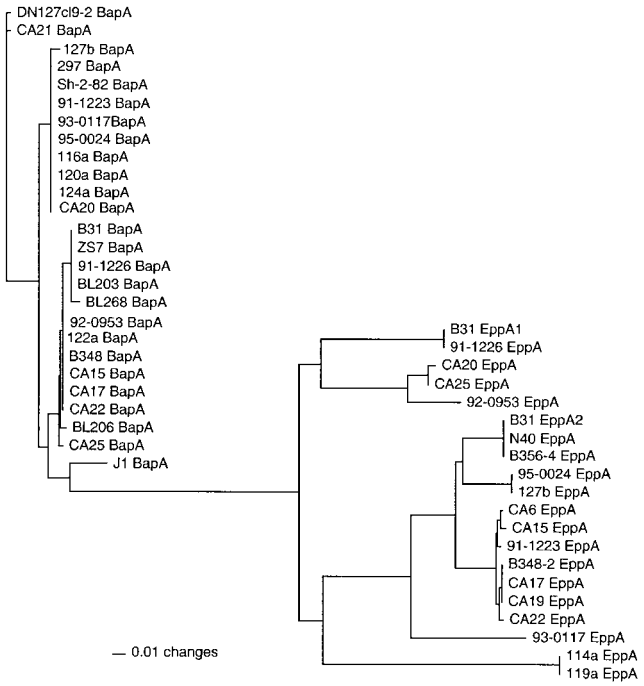

Fig. 2. Phylogenetic tree of BapA and EppA proteins. The partially determined BapA sequence of strain 90-1654, and those of strains encoding a truncated BapA protein, were not considered in this analysis. Only strains with completely determined EppA protein sequence were considered.

Four strains contained *bapA* genes that are identical to each other, but encode truncated proteins (Fig. 1A). All other *bapA* genes contain a run of seven Ts halfway through the gene, but the four variants contain one less T. If corrected for the frameshift, these genes would encode BapA proteins nearly identical in sequence to the Sh-2-82 BapA protein. Two of these strains, CA6 and CA19, were isolated from *I. pacificus* ticks collected in California, whereas the other two strains, 93-0107 and B356, were isolated from Lyme disease patients living in Wisconsin and New York, respectively (Table 1). Analyses of flanking DNA sequences demonstrated that at least two of these strains are only distantly related to each other (see below), indicating that the mutations probably arose independently. A number of other prokaryotes, including *Neisseria* and *Mycoplasma*, regulate expression of certain proteins by inactivating their genes through slipped-strand mispairing during replication of repetitive DNA sequences (Belland *et al.*, 1989; Murphy *et al.*, 1989; Yogev *et al.*, 1991). It has been hypothesized that slipped-strand mispairing modulates expression of the antigenic *B. burgdorferi* VraA lipoprotein (Labandeira-Rey *et al.*, 2001), and our results suggest that a similar regulatory process may also be at work in the *bapA* gene.

bapA genetic recombination events

Analysed strains of *B. burgdorferi* contain anywhere between six and ten different *erp* loci, but, at most, only one locus per strain contains an adjacent *bapA* gene (this work; Akins *et al.*, 1999; Casjens *et al.*, 2000). It is therefore most reasonable to assume that an ancestral *bapA* gene recombined into a region immediately 3′ of one or more *erp* locus at some point in the past. We examined several *bapA* loci for evidence of this postulated recombination event, and used physical linkage with distinctive plasmid markers to ascertain the genetic stability of *bapA*.

The examined *bapA* genes all contain a well-conserved 55 bp direct repeat sequence on either side of the gene (Fig. 3). Direct repeats are a hallmark of transposable elements, due to staggered strand-cutting during integration events (Shapiro, 1983). We note that the *bapA* gene does not bear significant homology to any known transposon, nor is it found elsewhere in the *B. burgdorferi* genome, as might be expected for a transposable element. However, consistent with the possibility that the ancestral *bapA* integrated into a cp32 via a transposition-like event, a single copy of the flanking sequence is found 3′ of some *B. burgdorferi erp* genes. For example, the strain N40 *ospF* gene lies adjacent to a nearly perfect copy of the *bapA* flanking sequence, while a somewhat divergent version is located 3′ of the strain B31 *erpK* gene. Other known *erp* loci contain either an extremely divergent version or an unrelated sequence (data not shown). No similar sequences are located adjacent to any known *eppA* gene. The regions 5′ of the *bapA* genes of several strains were examined, and all were found to contain a single *erp* gene. Comparison of these *erp* genes with those

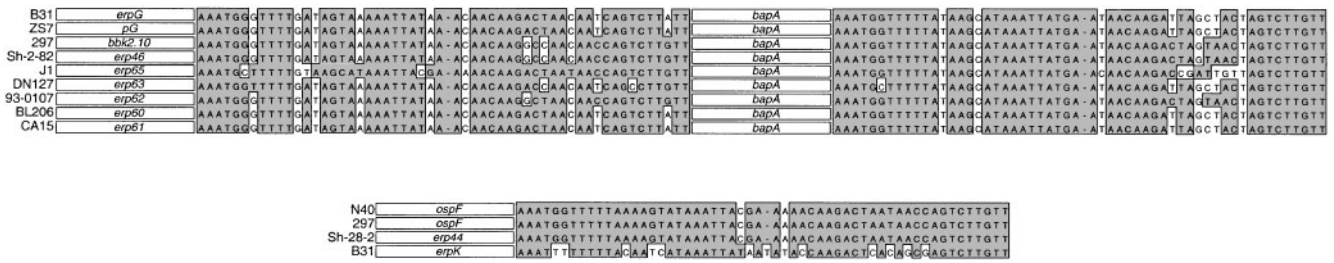

Fig. 3. Alignments of the 55 bp direct repeat elements flanking every known *bapA* gene, and homologous sequences located 3′ of certain *erp* genes. *bapA* and *erp* genes are indicated by white rectangles. Conserved nucleotides are boxed and shaded. Gaps introduced to maximize alignments are indicated by dashes.

of the well-characterized strains B31, 297 and N40 indicates that all *bapA*-linked *erp* genes are rather closely related and constitute a distinct clade (Fig. 4). These genetic similarities suggest that the current *bapA* loci all arose from a single insertion event into a region 3′ of just one ancestral *erp* gene.

Comparative analyses of *B. burgdorferi* cp32 plasmids indicate that recombination events between these plasmids periodically occur in nature (Akins *et al.*, 1999; Casjens *et al.*, 2000; Stevenson *et al.*, 1998b). Sequences located 3′ of the characterized *bapA* genes provide significant additional evidence of cp32 instability. Directly downstream of each 3′ direct repeat element is a conserved 8 bp sequence, followed by one of three very different DNA sequences (Fig. 5). Interestingly, the strain N40 *ospF* gene also contains the 8 bp conserved sequence, although the sequence beyond that point is unlike any adjacent to a studied *bapA* gene. Five of the examined strains, including 297, contain *bapA* genes followed by a copy of open reading frame 25 (ORF-25), which has no known homologue anywhere else in *B. burgdorferi* (Akins *et al.*, 1999). The *bapA* genes of four other strains, including B31, contain a short portion of ORF-25, followed by a gene of *B. burgdorferi* paralogue family 115, which is found on all cp32 plasmids (Amouriaux *et al.*, 1993; Casjens *et al.*, 1997, 2000; Stevenson *et al.*, 1997). The five strains with the intact ORF-25 also contain a paralogue family 115 gene, located approximately 1 kb downstream of ORF-25 (data not shown; Akins *et al.*, 1999). Strain J1 contains a completely different downstream sequence, a homologue of the BBM39 ORF that is located 3′ of the *erpK* gene of strain B31. Most likely, the B31-type flanking region arose through partial deletion of the ORF-25 of a 297-type plasmid, while the J1-type flanking region arose from an entirely different recombination event. Additionally, strain N40 is one of the few strains in our collection that lacks any evidence of a *bapA* gene. Most of the N40 *erp25* gene is very similar to the *bapA*-linked *erp* genes of strains 93-0107 and 297 (Table 1, Fig. 4), yet the 3′ end of the N40 *erp* gene differs greatly from those of these other two strains and is followed by a completely different downstream sequence (data not shown). These comparisons suggest that a relatively recent recombination event removed the *bapA* gene from that locus of N40.

The cp32 plasmids all contain a sequence-variable locus encoding proteins required for plasmid maintenance (Eggers *et al.*, 2002; Stevenson *et al.*, 2001). It is differences among these maintenance genes that enables compatibility of the many different cp32 family members carried by individual bacteria (Eggers *et al.*, 2002). Comparisons of those maintenance genes indicate high degrees of similarities among plasmids of different strains, which appear to correlate with incompatibility groups (Eggers *et al.*, 2002; Stevenson *et al.*, 1998b, 2001). To date, 12 such groups have

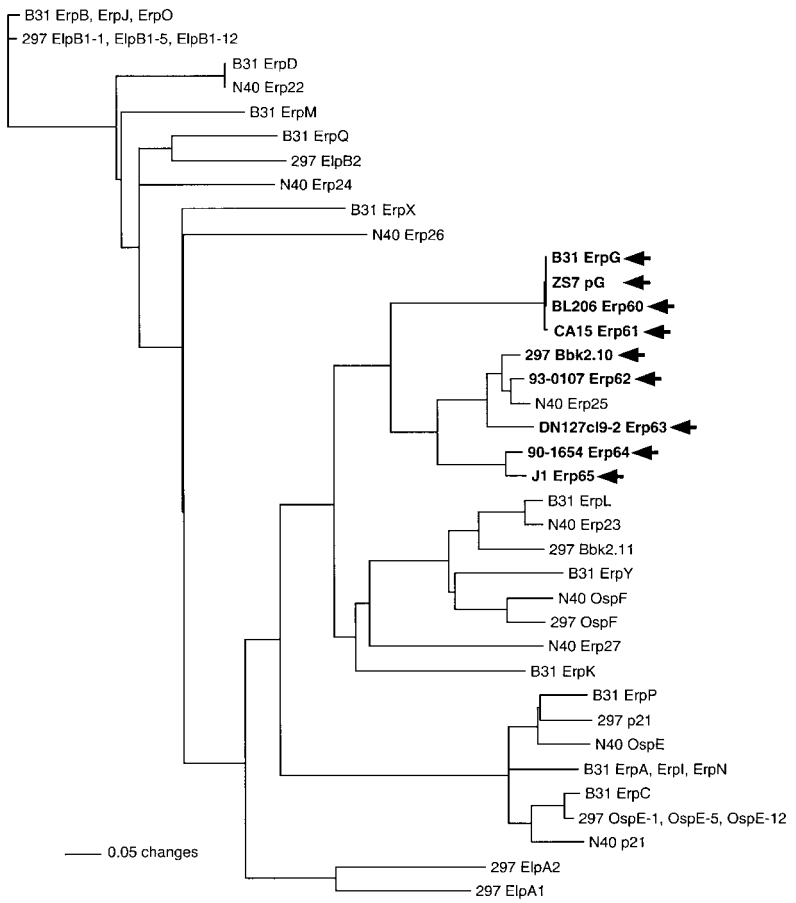

Fig. 4. Phylogenetic tree of Erp proteins from three well-characterized strains, B31, 297 and N40 (our unpublished results; Akins *et al.*, 1999; Casjens *et al.*, 2000; Lam *et al.*, 1994) and from strains characterized in the present study. Proteins encoded by loci that also contain a *bapA* gene are indicated in bold face type and by arrows.

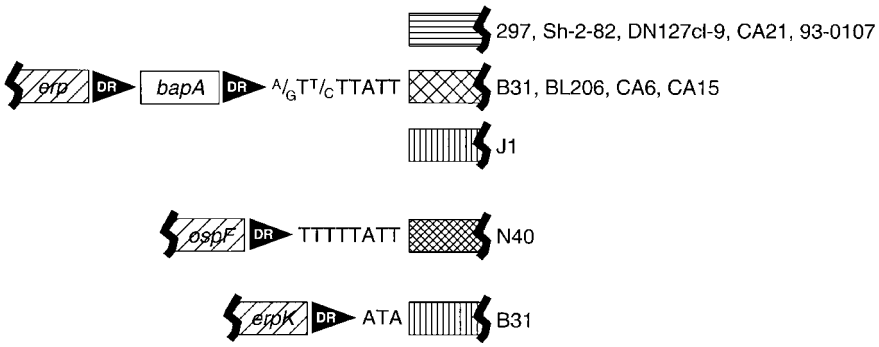

Fig. 5. Schematic of DNA sequences surrounding *bapA* genes. Direct repeat (DR) elements are indicated by arrowheads. An 8 bp conserved sequence is located immediately 3′ of the *bapA* downstream repeat element and 3′ of the strain N40 *ospF* gene. Three different types of DNA follow the *bapA*-linked conserved sequence, and strains in which each was identified are listed to the right. An unrelated fourth type of DNA sequence follows the strain N40 *ospF* conserved sequence. The DNA located 3′ of the strain J1 *bapA* gene is homologous to an ORF located immediately 3′ of the strain B31 *erpK* locus. The B31 *erpK* direct repeat element and this ORF are separated by a 3 bp sequence that does not correspond with the *bapA*- and *ospF*-associated conserved sequence.

been identified in the cp32 family (our unpublished results; Eggers *et al.*, 2002; Stevenson *et al.*, 1998b). The unique maintenance gene sequences of these different groups also provide physically linked genetic markers to evaluate recombination between different plasmids (Casjens *et al.*, 2000; Stevenson *et al.*, 1998b). In this manner, we determined that the BL206 *bapA* gene is located on a cp32-3 group plasmid, as is the strain B31 gene (data not shown; Stevenson *et al.*, 1998b). The 93-0107 *bapA* locus is carried on a cp32-5 group plasmid, the 297 and Sh-2-82 *bapA* genes are located on cp32-6 group plasmids, and the DN127cl9-2 and 90-1654 genes are on plasmids of the cp32-9 group. The CA15 *bapA* locus is carried on a type of cp32 never before observed, which we have designated group cp32-13. We conclude that past recombination events have shuffled *erp–bapA* loci between many different cp32 family members.

B. burgdorferi eppA genes

Bacteria were also examined for the second type of gene in paralogue family 95, *eppA*. Southern blot analysis using a probe derived from the strain B31 *eppA1* gene demonstrated hybridization with 63 % of the strains under low-stringency wash conditions, and 35 % under high-stringency conditions (Table 1). These blots indicated that all detected *eppA* genes were located on DNAs having electrophoretic mobilities comparable to 9 kb circular plasmids (data not shown). Next, PCR was performed using oligonucleotide primers complementary to sequences found in the strain B31 *eppA1* and *eppA2* genes (Table 2), yielding an *eppA* amplicon from 28 of 52 tested strains. An *eppA* amplicon could not be obtained from the uncloned B356 parental isolate, or from four out of the five B356 clones examined, but a gene was obtained for clone B356-4. Including those strains already known to contain an *eppA* gene, 38 of 54 (70 %) analysed *B. burgdorferi* isolates were demonstrated by either Southern blotting or PCR to contain an *eppA* gene, and 65 % contain both types of paralogue family 95 genes.

Nineteen *eppA* PCR products were chosen at random and sequenced. Alignment of the predicted EppA proteins indicated a great deal of both identity and diversity among the examined isolates, with amino acid sequence identities that ranged from 54 to 100 % (Fig. 1B, Fig. 2). Most predicted EppA proteins have molecular masses of approximately 21 kDa and basic pI values ranging between 8·6 and 9·7. Curiously, the EppA proteins of strains 114a and 119a are predicted to have sizes similar to all other analysed EppA proteins, yet have very acidic pI values of approximately 4·9. There was considerable evidence of past insertion or deletion mutations having occurred within these genes, although high degrees of stability were also observed. The *eppA* gene of strain 91-1226, isolated from a human patient in Maryland, is identical to the *eppA1* allele of strain B31. The EppA proteins of the New York strains N40 and B356-4 are predicted to be identical to the B31 EppA2 protein. The *eppA* genes of the New York patient isolates 114a and 119a exhibited a very high degree of divergence from other genes, but were identical to each other. Two bacteria isolated from ticks collected in California, strains CA17 and CA19, contain e*ppA* genes identical to that of strain B348, which was isolated from a human in New York. Only one of the *eppA* genes is predicted to encode a protein with the RGD motif found in most BapA proteins (see above), although all contain a variation on that amino acid sequence (Fig. 1).

As noted above, all *eppA* genes appeared to be located on approximately 9 kb circular plasmids. The complete sequences of three related 9 kb plasmids from three different *B. burgdorferi* strains have been determined, and found to exhibit limited sequence conservation even in homologous loci (Dunn *et al.*, 1994; Fraser *et al.*, 1997; Stewart *et al.*, 2001). These variations are probably reflective of the

tortuous evolutionary paths these plasmids have followed in their descent from cp32 plasmids (Casjens *et al.*, 2000). Due to an inability to identify informative, conserved sequences in the previously sequenced plasmids that would be suitable for PCR analysis, we did not attempt to characterize DNAs adjacent to the newly identified *eppA* genes.

The widespread variability in EppA sequence suggests that it is under some sort of selective pressure. EppA is proposed to be an integral outer-membrane protein (Champion *et al.*, 1994) and may therefore be exposed to the host immune system. However, we note that the examined bacteria were isolated from a wide variety of hosts, including humans, mice, birds, and ticks that feed on a wide variety of vertebrates. It has been proposed that amino acid sequence differences of some proteins between strains influence the efficiency of bacteria for infection of different host species (Stevenson *et al.*, 2001). As one example, sequence differences among Erp proteins alter their ability to bind complement factor H, which is thought to contribute to the host range of Lyme disease borreliae (Kraiczy *et al.*, 2001b; Kurtenbach *et al.*, 2002a, b; Stevenson *et al.*, 2002). Perhaps EppA proteins also interact with host proteins that differ in sequence between diverse hosts in nature? Continued studies of this bacterial protein and elucidation of its function(s) will resolve this question.

Although significant diversity was noted among the *eppA* genes of many strains, there were also several examples of identity between strains. The *eppA* genes of Wisconsin isolate 95-0024 and New York isolate 127b are identical, as are those of the New York isolate B348-2 and California strains CA17, CA19 and CA22. It has been proposed that *B. burgdorferi* is largely a clonal organism (Dykhuizen & Baranton, 2001; Qiu *et al.*, 2002), so the occurrence of similar sequences in bacteria from widely separated geographical regions may be a reflection of that characteristic. However, genetic exchange between bacteria may also be responsible for the widespread nature of certain alleles. For example, strains B31 and N40 have identical *eppA* genes but differ at many other examined loci (our unpublished results; Liang *et al.*, 2002; Roberts *et al.*, 1998; Stevenson & Barthold, 1994). These data argue against a hypothesis that only very small fragments of DNA can be transferred between *B. burgdorferi* (Dykhuizen & Baranton, 2001). Further comparisons of strains possessing identical *eppA* and other loci will continue to test of the validity of that hypothesis.

Analysis of infected human sera

Serum samples obtained from 49 Lyme disease patients and six healthy humans were assayed for the presence of antibodies that recognized recombinant B31 BapA, EppA1, and EppA2 proteins (Table 3). Six out of 49 (12 %) Lyme disease patient sera contained antibodies that recognized the B31 BapA protein, confirming that *B. burgdorferi* produces this protein during mammalian infection. Of the 16 strains isolated from patients who also provided serum samples, 14 contained a *bapA* gene, yet only one of these patients produced detectable levels of antibodies against BapA. As described above, all of those infectious isolates encoded BapA proteins sharing over 95 % amino acid identities with the B31 protein. These results indicate that the failure of some patients to produce antibodies that recognize the B31 BapA fusion was not because they were infected by bacteria that lacked a gene similar to that of strain B31.

Eighteen out of 49 (37 %) patient serum samples contained antibodies that recognized the EppA1 fusion protein, and 17 out of 49 (35 %) contained antibodies that bound the EppA2 fusion protein. A positive correlation was noted between the production of EppA-binding antibodies and the presence of an *eppA* gene in the infecting bacteria: of the bacteria isolated from patients who also provided serum samples, eight strains contained *eppA* genes and seven of those patients were seropositive. Ten patients produced antibodies that recognized both EppA1 and EppA2, indicating that these two proteins contain some similar epitopes. However, some regions of the two B31 proteins studied are antigenically distinct, since several patients produced antibodies that recognized only one recombinant protein. This raises the possibility that many of the serologically negative Lyme disease patients actually produced antibodies to the EppA proteins of their infecting organisms, which contained epitopes that differ from those of the strain B31 EppA1 and EppA2 proteins. The diversity of *eppA* sequences among the strains analysed in this study suggests the possibility of an even wider degree of variation throughout nature, some of which is too great to have been detected by our PCR primers or Southern blot probes.

Conclusions

A large majority of Lyme disease spirochaetes carry a member of gene paralogue family 95, with most bacteria examined containing both a *bapA* and an *eppA* gene. These high proportions strongly suggest that the ability to produce BapA and/or EppA confers a selective advantage to *B. burgdorferi* in nature. A high degree of sequence conservation was noted among *bapA* genes, although it was also evident that DNAs adjacent to these genes have undergone significant deletion and replacement mutations. The plasmid location of *bapA* genes is also quite variable, with analysed loci being mapped to five different cp32 plasmid groups. The stability of the *bapA* gene in the face of such extensive genetic rearrangements suggests that the encoded protein performs a function that does not permit much structural variation. On the other hand, mutations have run riot through the *eppA* locus. Extensive variation is also evident throughout the 9 kb circular plasmids that carry *eppA* genes, and these two phenomena may be linked. Finally, the highly conserved BapA does not appear to be highly antigenic, while the extremely variable EppA was much more antigenic. The reasons for these genetic and antigenic variations are presently unknown, and indicate a need for continued studies of these intriguing genes and

their proteins to determine their roles in *B. burgdorferi* biology and the pathogenesis of Lyme disease.

ACKNOWLEDGEMENTS

This study was funded by US Public Health Services grant RO1-AI44254 to Brian Stevenson. Jennifer Miller is supported by a University of Kentucky Dissertation Fellowship and US Public Health Services Training Grant 5T32-AI49795.

We thank Sherwood Casjens, Russell Johnson, Martin Schriefer, Tom Schwan, Ira Schwartz and Gary Wormser for providing *B. burgdorferi* strains and Lyme disease patient serum samples; and Kelly Babb, Nazira El-Hage, Melissa Hines, Julie Stewart and Natalie Mickelsen for technical assistance and thoughtful discussions.

REFERENCES

Akins, D. R., Porcella, S. F., Popova, T. G., Shevchenko, D., Baker, S. I., Li, M., Norgard, M. V. & Radolf, J. D. (1995). Evidence for in vivo but not in vitro expression of a *Borrelia burgdorferi* outer surface protein F (OspF) homologue. *Mol Microbiol* **18**, 507–520.

Akins, D. R., Caimano, M. J., Yang, X., Cerna, F., Norgard, M. V. & Radolf, J. D. (1999). Molecular and evolutionary analysis of *Borrelia burgdorferi* 297 circular plasmid-encoded lipoproteins with OspE- and OspF-like leader peptides. *Infect Immun* **67**, 1526–1532.

Alitalo, A., Meri, T., Lankinen, H., Seppälä, I., Lahdenne, P., Hefty, P. S., Akins, D. & Meri, S. (2002). Complement inhibitor factor H binding to Lyme disease spirochetes is mediated by inducible expression of multiple plasmid-encoded outer surface protein E paralogs. *J Immunol* **169**, 3847–3853.

Amouriaux, P., Assous, M., Margarita, D., Baranton, G. & Saint Girons, I. (1993). Polymerase chain reaction with the 30-kb circular plasmid of *Borrelia burgdorferi* B31 as a target for detection of the Lyme borreliosis agents in cerebrospinal fluid. *Res Microbiol* **144**, 211–219.

Baranton, G., Postic, D., Saint Girons, I., Boerlin, P., Piffaretti, J.-C., Assous, M. & Grimont, P. A. D. (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.

Barbour, A. G. (1984). Isolation and cultivation of Lyme disease spirochetes. *Yale J Biol Med* **57**, 521–525.

Barthold, S. W., Moody, K. D., Terwilliger, G. A., Duray, P. H., Jacoby, R. O. & Steere, A. C. (1988). Experimental Lyme arthritis in rats infected with *Borrelia burgdorferi*. *J Infect Dis* **157**, 842–846.

Bauer, Y., Hofmann, H., Jahraus, O., Mytilineos, J., Simon, M. M. & Wallich, R. (2001). Prominent T cell response to a selectively *in vivo* expressed *Borrelia burgdorferi* outer surface protein (pG) in patients with Lyme disease. *Eur J Immunol* **31**, 767–776.

Belland, R. J., Morrison, S. G., van der Ley, P. & Swanson, J. (1989). Expression and phase variation of gonnococcal P.II genes in *Escherichia coli* involves ribosomal frameshifting and slipped-strand mispairing. *Mol Microbiol* **3**, 777–786.

Bissett, M. L. & Hill, W. (1987). Characterization of *Borrelia burgdorferi* strains isolated from *Ixodes pacificus* ticks in California. *J Clin Microbiol* **25**, 2296–2301.

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.

Casjens, S., van Vugt, R., Tilly, K., Rosa, P. A. & Stevenson, B. (1997). Homology throughout the multiple 32-kilobase circular plasmids present in Lyme disease spirochetes. *J Bacteriol* **179**, 217–227.

Casjens, S., Palmer, N., van Vugt, R. & 12 other authors (2000). A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs of an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Mol Microbiol* **35**, 490–516.

Champion, C. I., Blanco, D. R., Skare, J. T., Haake, D. A., Giladi, M., Foley, D., Miller, J. N. & Lovett, M. A. (1994). A 9·0-kilobase-pair circular plasmid of *Borrelia burgdorferi* encodes an exported protein: evidence for expression only during infection. *Infect Immun* **62**, 2653–2661.

Cinco, M., Banfi, E., Balanzin, D., Caccio, S., Graziosi, G. & Fattorini, P. (1989). Restriction endonuclease analysis of four *Borrelia burgdorferi* strains. *FEMS Microbiol Immunol* **1**, 511–514.

Dunn, J. J., Buchstein, S. R., Butler, L.-L., Fisenne, S., Polin, D. S., Lade, B. N. & Luft, B. J. (1994). Complete nucleotide sequence of a circular plasmid from the Lyme disease spirochete, *Borrelia burgdorferi*. *J Bacteriol* **176**, 2706–2717.

Dykhuizen, D. E. & Baranton, G. (2001). Implications of a low rate of horizontal transfer in *Borrelia*. *Trends Microbiol* **9**, 344–350.

Eberhard, W. G. (1989). Why do bacterial plasmids carry some genes and not others? *Plasmid* **21**, 167–174.

Eggers, C. H., Caimano, M. J., Clawson, M. L., Miller, W. G., Samuels, D. S. & Radolf, J. D. (2002). Identification of loci critical for replication and compatibility of a *Borrelia burgdorferi* cp32 plasmid and use of a cp32-based shuttle vector for the expression of fluorescent reporters in the Lyme disease spirochaete. *Mol Microbiol* **43**, 281–295.

El-Hage, N., Lieto, L. D. & Stevenson, B. (1999). Stability of *erp* loci during *Borrelia burgdorferi* infection: recombination is not required for chronic infection of immunocompetent mice. *Infect Immun* **67**, 3146–3150.

El-Hage, N., Babb, K., Carroll, J. A., Lindstrom, N., Fischer, E. R., Miller, J. C., Gilmore, R. D., Jr, Mbow, M. L. & Stevenson, B. (2001). Surface exposure and protease insensitivity of *Borrelia burgdorferi* Erp (OspEF-related) lipoproteins. *Microbiology* **147**, 821–830.

Fraser, C. M., Casjens, S., Huang, W. M. & 35 other authors (1997). Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **390**, 580–586.

Hefty, P. S., Jolliff, S. E., Caimano, M. J., Wikel, S. K., Radolf, J. D. & Akins, D. R. (2001). Regulation of OspE-related, OspF-related, and Elp lipoproteins of *Borrelia burgdorferi* strain 297 by mammalian host-specific signals. *Infect Immun* **69**, 3618–3627.

Hellwage, J., Meri, T., Heikkilä, T., Alitalo, A., Panelius, J., Lahdenne, P., Seppälä, I. J. T. & Meri, S. (2001). The complement regulatory factor H binds to the surface protein OspE of *Borrelia burgdorferi*. *J Biol Chem* **276**, 8427–8435.

Hughes, C. A. N., Kodner, C. B. & Johnson, R. C. (1992). DNA analysis of *Borrelia burgdorferi* NCH-1, the first northcentral U.S. human Lyme disease isolate. *J Clin Microbiol* **30**, 698–703.

Hyde, F. W. & Johnson, R. C. (1988). Characterization of a circular plasmid from *Borrelia burgdorferi*, etiologic agent of Lyme disease. *J Clin Microbiol* **26**, 2203–2205.

Hynes, R. O. (1992). Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* **69**, 11–25.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. *Trends Biochem Sci* **23**, 403–405.

Kraiczy, P., Skerka, C., Brade, V. & Zipfel, P. F. (2001a). Further characterization of complement regulator-acquiring surface proteins of *Borrelia burgdorferi*. *Infect Immun* **69**, 7800–7809.

Kraiczy, P., Skerka, C., Kirschfink, M., Brade, V. & Zipfel, P. F. (2001b). Immune evasion of *Borrelia burgdorferi* by acquisition of human complement regulators FHL-1/reconectin and factor H. *Eur J Immunol* **31**, 1674–1684.

Kryuchechnikov, V. N., Korenberg, E. I., Shcherbakov, S. V., Kovalevsky, Y. V. & Levin, M. L. (1988). Identification of *Borrelia* isolated in the USSR from *Ixodes persulcatus* Schulze ticks. *J Microbiol Epidemiol Immunobiol* **12**, 41–44.

Kurtenbach, K., DeMichelis, S., Etti, S., Schäfer, S. M., Sewell, H.-S., Brade, V. & Kraiczy, P. (2002a). Host association of *Borrelia burgdorferi* sensu lato - the key role of host complement. *Trends Microbiol* **10**, 74–79.

Kurtenbach, K., Schäfer, S. M., Sewell, H.-S., Peacey, M., Hoodless, A., Nuttall, P. A. & Randolph, S. E. (2002b). Differential survival of Lyme borreliosis spirochetes in ticks that feed on birds. *Infect Immun* **70**, 5893–5895.

Kurtti, T. J., Munderloh, U. G., Johnson, R. C. & Ahlstrand, G. G. (1987). Colony formation and morphology in *Borrelia burgdorferi*. *J Clin Microbiol* **25**, 2054–2058.

Labandeira-Rey, M. & Skare, J. T. (2001). Decreased infectivity in *Borrelia burgdorferi* strain B31 is associated with loss of linear plasmid 25 or 28-1. *Infect Immun* **69**, 446–455.

Labandeira-Rey, M., Baker, E. A. & Skare, J. T. (2001). VraA (BBI16) protein of *Borrelia burgdorferi* is a surface-exposed antigen with a repetitive motif that confers partial protection against experimental Lyme borreliosis. *Infect Immun* **69**, 1409–1419.

Lam, T. T., Nguyen, T.-P. K., Montgomery, R. R., Kantor, F. S., Fikrig, E. & Flavell, R. A. (1994). Outer surface proteins E and F of *Borrelia burgdorferi*, the agent of Lyme disease. *Infect Immun* **62**, 290–298.

Liang, F. T., Nelson, F. K. & Fikrig, E. (2002). DNA microarray assessment of putative *Borrelia burgdorferi* lipoprotein genes. *Infect Immun* **70**, 3300–3303.

Marconi, R. T., Konkel, M. E. & Garon, C. F. (1993). Variability of *osp* genes and gene products among species of Lyme disease spirochetes. *Infect Immun* **61**, 2611–2617.

Masuzawa, T., Okada, Y., Beppu, Y., Oku, T., Kawamori, F. & Yanagihara, Y. (1991). Immunological properties of *Borrelia burgdorferi* isolated from the *Ixodes ovatus* in Shizuoka, Japan. *Microbiol Immunol* **35**, 913–919.

McLean, R. G., Ubico, S. R., Hughes, C. A., Engstrom, S. M. & Johnson, R. C. (1993). Isolation and characterization of *Borrelia burgdorferi* from blood of a bird captured in the Saint Croix River Valley. *J Clin Microbiol* **31**, 2038–2043.

Miller, J. C., Bono, J. L., Babb, K., El-Hage, N., Casjens, S. & Stevenson, B. (2000a). A second allele of *eppA* in *Borrelia burgdorferi* strain B31 is located on the previously undetected circular plasmid cp9-2. *J Bacteriol* **182**, 6254–6258.

Miller, J. C., El-Hage, N., Babb, K. & Stevenson, B. (2000b). *Borrelia burgdorferi* B31 Erp proteins that are dominant immunoblot antigens of animals infected with isolate B31 are recognized by only a subset of human Lyme disease patient sera. *J Clin Microbiol* **38**, 1569–1574.

Murphy, G. L., Connell, T. D., Barritt, D. S., Koomey, M. & Cannon, J. G. (1989). Phase variation of gonnococcal protein II: regulation of gene expression by slipped strand mispairing of a repetitive DNA sequence. *Cell* **56**, 539–547.

Nohlmans, L., de Boer, R., van den Bogaard, A. & van Boven, C. (1995). Genotypic and phenotypic analysis of *Borrelia burgdorferi* isolates from the Netherlands. *J Clin Microbiol* **33**, 119–125.

Postic, D., Assous, M. V., Grimont, P. A. D. & Baranton, G. (1994). Diversity of *Borrelia burgdorferi* sensu lato evidenced by restriction fragment length polymorphism of *rrf* (5S)-*rrl* (23S) intergenic spacer amplicons. *Int J Syst Bacteriol* **44**, 743–752.

Purser, J. E. & Norris, S. J. (2000). Correlation between plasmid content and infectivity in *Borrelia burgdorferi*. *Proc Natl Acad Sci U S A* **97**, 13865–13870.

Qiu, W.-G., Dykhuizen, D. E., Acosta, M. S. & Luft, B. J. (2002). Geographic uniformity of the Lyme disease spirochete (*Borrelia burgdorferi*) and its shared history with tick vector (*Ixodes scapularis*) in the northeastern United States. *Genetics* **160**, 833–849.

Roberts, W. C., Mullikin, B. A., Lathigra, R. & Hanson, M. S. (1998). Molecular analysis of sequence heterogeneity among genes encoding decorin binding proteins A and B of *Borrelia burgdorferi* sensu lato. *Infect Immun* **66**, 5275–5285.

Rosa, P., Samuels, D. S., Hogan, D., Stevenson, B., Casjens, S. & Tilly, K. (1996). Directed insertion of a selectable marker into a circular plasmid of *Borrelia burgdorferi*. *J Bacteriol* **178**, 5946–5953.

Schwan, T. G., Burgdorfer, W. & Garon, C. F. (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.

Schwan, T. G., Schrumpf, M. E., Karstens, R. H., Clover, J. R., Wong, J., Daugherty, M., Struthers, M. & Rosa, P. A. (1993). Distribution and molecular analysis of Lyme disease spirochetes, *Borrelia burgdorferi*, isolated from ticks throughout California. *J Clin Microbiol* **31**, 3096–3108.

Shapiro, J. A. (1983). *Mobile Genetic Elements*. New York: Academic Press.

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*. *J Clin Microbiol* **28**, 1329–1337.

Steere, A. C., Grodzicki, R. L., Craft, J. E., Shrestha, M., Kornblatt, A. N. & Malawista, S. E. (1984). Recovery of Lyme disease spirochetes from patients. *Yale J Biol Med* **57**, 557–560.

Stevenson, B. & Barthold, S. W. (1994). Expression and sequence of outer surface protein C among North American isolates of *Borrelia burgdorferi*. *FEMS Microbiol Lett* **124**, 367–372.

Stevenson, B., Tilly, K. & Rosa, P. A. (1996). A family of genes located on four separate 32-kilobase circular plasmids in *Borrelia burgdorferi* B31. *J Bacteriol* **178**, 3508–3516.

Stevenson, B., Casjens, S., van Vugt, R., Porcella, S. F., Tilly, K., Bono, J. L. & Rosa, P. (1997). Characterization of cp18, a naturally truncated member of the cp32 family of *Borrelia burgdorferi* plasmids. *J Bacteriol* **179**, 4285–4291.

Stevenson, B., Bono, J. L., Schwan, T. G. & Rosa, P. (1998a). *Borrelia burgdorferi* Erp proteins are immunogenic in mammals infected by tick bite, and their synthesis is inducible in cultured bacteria. *Infect Immun* **66**, 2648–2654.

Stevenson, B., Casjens, S. & Rosa, P. (1998b). Evidence of past recombination events among the genes encoding the Erp antigens of *Borrelia burgdorferi*. *Microbiology* **144**, 1869–1879.

Stevenson, B., Zückert, W. R. & Akins, D. R. (2001). Repetition, conservation, and variation: the multiple cp32 plasmids of *Borrelia* species. In *The Spirochetes: Molecular and Cellular Biology*, pp. 87–100. Edited by M. H. Saier & J. García-Lara. Oxford: Horizon Press.

Stevenson, B., El-Hage, N., Hines, M. A., Miller, J. C. & Babb, K. (2002). Differential binding of host complement inhibitor factor H by *Borrelia burgdorferi* Erp surface proteins: a possible mechanism underlying the expansive host range of Lyme disease spirochetes. *Infect Immun* **70**, 491–497.

Stewart, P. E., Thalken, R., Bono, J. L. & Rosa, P. (2001). Isolation of a circular plasmid region sufficient for autonomous replication and transformation of infectious *Borrelia burgdorferi*. *Mol Microbiol* **39**, 714–721.

Swofford, D. L. (2000). *PAUP*, Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4. Sunderland, MA: Sinauer Associates.

Wallich, R., Brenner, C., Kramer, M. D. & Simon, M. M. (1995). Molecular cloning and immunological characterization of a novel linear-plasmid-encoded gene, *pG*, of *Borrelia burgdorferi* expressed only in vivo. *Infect Immun* **63**, 3327–3335.

Wang, G., Ojaimi, C., Iyer, R., Saksenberg, V., McClain, S. A., Wormser, G. P. & Schwartz, I. A. (2001). Impact of genotypic variation of *Borrelia burgdorferi* sensu stricto on kinetics of dissemination and severity of disease in C3H/HeJ mice. *Infect Immun* **69**, 4303–4312.

Xu, Y. & Johnson, R. C. (1995). Analysis and comparison of plasmid profiles of *Borrelia burgdorferi* sensu lato strains. *J Clin Microbiol* **33**, 2679–2685.

Xu, Y., Kodner, C., Coleman, L. & Johnson, R. C. (1996). Correlation of plasmids with infectivity of *Borrelia burgdorferi* sensu stricto type strain B31. *Infect Immun* **64**, 3870–3876.

Yogev, D., Rosengarten, R., Watson-McKown, R. & Wise, K. S. (1991). Molecular basis of *Mycoplasma* surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5′ regulatory sequences. *EMBO J* **10**, 4069–4079.

Zimmer, G., Schaible, U. E., Kramer, M. D., Mall, G., Museteanu, C. & Simon, M. M. (1990). Lyme carditis in immunodeficient mice during experimental infection of *Borrelia burgdorferi*. *Virchows Arch Pathol Ana Histopathol* **417**, 129–135.