JOURNAL OF BACTERIOLOGY, Sept. 1998, p. 4974–4981
0021-9193/98/$04.00+0

Cloning and Molecular Characterization of a Multicopy, Linear Plasmid-Carried, Repeat Motif-Containing Gene from *Borrelia turicatae*, a Causative Agent of Relapsing Fever

JASON A. CARLYON AND RICHARD T. MARCONI*

Department of Microbiology and Immunology, Medical College of Virginia at Virginia Commonwealth University, Richmond, Virginia 23298-0678

Received 6 May 1998/Accepted 22 July 1998

***Borrelia turicatae* is one of several spirochete species that can cause relapsing fever. Here, we describe the identification and characterization of a gene from *B. turicatae* and other relapsing-fever spirochetes that exhibits homology with the *rep*$^+$ and ORF-E gene families of the Lyme disease spirochetes. This gene, which we have designated *repA*, encodes a putative protein of 30.2 kDa with an isoelectric point of 4.69. The central region of RepA harbors a series of amino acid repeat motifs which exhibit homology with casein kinase 2 phosphorylation sites. Through Southern hybridization analyses, we demonstrate that *repA* (or a closely related sequence) is multicopy in the relapsing-fever spirochetes and is carried on variably sized linear plasmids in both *Borrelia parkeri* and *B. turicatae*. Transcriptional analyses demonstrate that *repA* is expressed, albeit at low levels, during in vitro cultivation of *B. turicatae*. Transcriptional start site analysis revealed that *repA* is preceded by a consensus ribosomal binding site and an appropriately spaced promoter element. The sequence conservation, unique features, and multicopy status of *repA* and its homologs suggest that RepA may play an important genus-wide role in the biology of the *Borrelia*.**

Borrelia turicatae and the closely related spirochetes *Borrelia hermsii* and *Borrelia parkeri* are causative agents of New World tick-borne relapsing fever. In North America, relapsing fever occurs primarily in the mountains and semiarid plains of the western United States and Mexico (2). Tick-borne relapsing fever is a zoonotic disease transmissable to humans through the bite of infected argasid and *Ornithodoros* ticks (11). In addition to the importance of *B. turicatae* and the relapsing fever spirochetes as causative agents of human disease, the study of these pathogens is important for other reasons as well. The causative agents of relapsing fever and Lyme disease share numerous genetic and physiological traits. One of the more striking shared features of the *Borrelia* species is their genome, which is comprised of variably sized linear and circular DNA species (3). In addition, the relapsing-fever spirochetes also carry homologs of some Lyme disease spirochete genes such as *ospC* (19), which is thought to be important in the transmission of the spirochetes from ticks to mammals (24). The *ospC* homolog in *B. hermsii*, *vmp33*, has been demonstrated to be a member of the *vmp* gene family (6). The existence of homologs of known virulence factors in multiple *Borrelia* species may indicate that these factors play a genus-wide role in *Borrelia* pathogenesis. In light of the similarities among *Borrelia* species, it has been suggested that *B. turicatae* could serve as a model organism for studying the molecular pathogenesis of *Borrelia* infections (21). One advantage in utilizing *B. turicatae* for this purpose is that unlike the Lyme disease spirochetes, this organism reaches high population densities in experimental animal models. This would greatly facilitate studies designed to assess the in vivo expression and function of putative virulence factors.

* Corresponding author. Mailing address: Department of Microbiology and Immunology, Medical College of Virginia at Virginia Commonwealth University, Richmond, VA 23298-0678. Phone: (804) 828-3779. Fax: (804) 828-9946. E-mail: RMARCONI@hsc.vcu.edu.

Cloning and characterization of a homolog of the Lyme disease spirochete *rep*$^+$ gene family in the relapsing-fever spirochete *B. turicatae*. In this study, we cloned and characterized from *B. turicatae* a homolog of the Lyme disease spirochete *rep*$^+$ gene family, which we have designated *repA* (all isolates analyzed in this report are described in Table 1). The recombinant clone carrying this gene (pBt2.2) was serendipitously recovered in the course of screening a *B. turicatae* (isolate 91E135) genomic library for homologs of the *Borrelia burgdorferi* sensu lato complex upstream homology box gene family (20). The plasmid library was generated by the ligation of *Xba*I-digested *B. turicatae* 91E135 genomic DNA into the polylinker region of pGem-3Zf(+) (Promega). Ligation products were transformed into high-efficiency JM109 cells (Promega) and plated onto Luria-Bertani plates containing ampicillin (100 μg ml^{-1}), IPTG (isopropyl-β-D-thiogalactopyranoside) (0.5 mM), and X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (80 μg ml^{-1}). Colony lifts were performed with Genescreen membranes as described by the manufacturer (DuPont-NEN). Cycle sequencing with the Sequitherm Excel DNA Sequencing Kit (Epicentre Technologies) of recombinant clone pBt2.2 revealed that it carries a 2.2-kb insert containing a 786-bp open reading frame (ORF). The ORF encodes a putative protein of 262 amino acids (Fig. 1) with a predicted molecular mass of 30.2 kDa. Its deduced pI of 4.69 is unusual in that the median deduced pI of *B. burgdorferi* proteins is 9.7 (12). A striking feature of this ORF is its central region (nucleotides 212 to 583), which harbors a series of direct repeat elements comprising 47.7% of the ORF. The longest repeat is 28 amino acids in length (Fig. 1). Ten smaller imperfect repeat units of 11 amino acids are also present. The tripeptide KID is present in each of these 10 copies and is perfectly conserved except in the last repeat copy, where it is KIE. Five of the KID(E) motifs are preceded by either T or S. This sequence is identical to that of the putative casein kinase 2 phosphorylation motif, S/T-X-X-acidic amino acid (DNE or Q) (22, 26). The repeat motif-carrying region of the protein is predicted to have a positive Jameson-Wolf antigenic index (13), but Chou-Fasman surface exposure probability predictions (9) for this as well as other regions of the protein are low. Interestingly, the carboxy terminus of the protein is predicted to be the most hydrophobic portion of the protein, with a Kyte and Doolittle hydrophobicity index (14) greater than 1.3. This may suggest that this region of the protein is embedded in the membrane. Since RepA lacks any obvious export signals, it is possible that it may associate with the inner leaflet of the inner membrane. Computer analyses with the Motifs program (Genetics Computer Group) did not reveal the presence of other amino acid sequence motifs that might give additional clues as to the cellular function or location of RepA.

Similarity between the ORF carried by pBt2.2 and several previously characterized Lyme disease spirochete genes of unknown function was revealed through a gapped BLAST search of the databases. Genes exhibiting homology include the *rep*$^+$ (23) and ORF-E gene families (28, 29), BBG33 (12), BBF03 (12), and *p21* (note that the *p21* gene in question [26] is different from the upstream homology box-flanked *p21* gene de-

TABLE 1. *Borrelia* isolates used in this study[a]

Isolate	Origin	Additional information and associated disease
Borrelia turicatae 91E135	*Ornithodoros turicatae*, United States	Low passage; infectious; relapsing fever
Borrelia turicatae OZ-1	Clone of 91E135 (4)	Low passage; infectious; relapsing fever
Borrelia parkeri	*Ornithodoros parkeri*, United States	High passage; infectious; relapsing fever
Borrelia hermsii HS-1	*Ornithodoros hermsii*, United States	Low passage; infectious; relapsing fever
Borrelia hermsii Man	Human blood isolate, United States	Low passage; infectious; relapsing fever
Borrelia hermsii Yor-1	Experimentally infected mouse	This isolate originated from a human relapsing fever patient in the United States; low passage; infectious
Borrelia japonica IKA-2	*Ixodes ovatus*, Japan	Low passage; infectivity unknown; Lyme disease
Borrelia anserina Ba-2	Avian isolate, United States	High passage; infectivity unknown; avian pathogen
Borrelia coriaceae CO53	*Ornithodoros coriaceus*, United States	Low passage; infectivity unknown; putative causative agent of epizootic bovine abortion

[a] All isolates were cultivated in standard BSK-H medium (Sigma) supplemented with 6% rabbit serum. For the relapsing-fever spirochetes, rabbit serum was added to a final concentration of 12%.

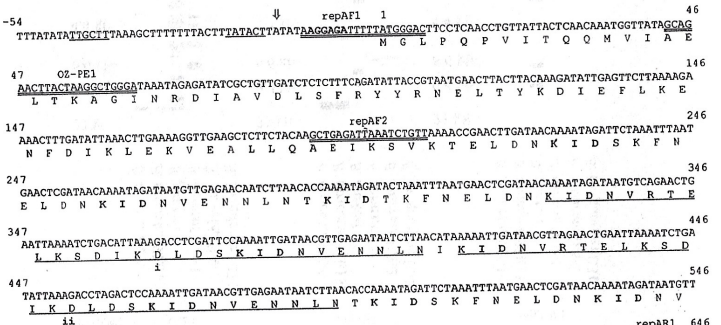

FIG. 1. Nucleotide and amino acid sequence of the *B. turicatae* 91E135 *repA* gene. The figure lists the sequence determined from the pBt2.2 clone. The binding sites of all probes and primers used in this study are indicated by double underlines. The actual sequences of the OZI-PEI and repAR1 primers are the inverse complements of those underlined. The largest of the amino acid repeats (28 amino acids) present in RepA is indicated by a single underline, and the two copies are denoted by i and ii. The tripeptide repeat motif KID(E) is indicated by bold lettering. The transcriptional start site identified in this report is indicated by an arrow, the ribosomal binding site by bold lettering, and the −10 and −35 elements by underlining.

TABLE 2. Amino acid similarity and identity values of *repA* of *B. turicatae* and *repA*-related genes of *B. burgdorferi* sensu lato isolates[a]

Protein or ORF	Amino acid similarity or identity (%)											
	RepA	BBF03	BBG33	ORF-E (cp18)	ORF-E (cp32)	ORF-E (lp50)	P21	Rep*-1	Rep*-2	Rep*-3	Rep*-4	Rep*-5
RepA		40.12	46.31	39.89	54.68	58.43	42.19	66.58	61.03	68.11	66.33	66.01
BBF03	33.14		62.79	40.61	42.77	46.71	60.12	54.32	57.23	52.07	56.98	56.98
BBG33	39.75	58.72		42.53	49.75	48.04	54.54	69.40	62.62	71.20	64.85	70.30
ORF-E (cp18)	32.45	29.70	35.11		62.57	66.67	36.90	38.33	53.68	39.44	45.46	47.22
ORF-E (cp32)	36.95	34.34	41.79	58.82		86.03	42.63	68.64	61.14	61.91	68.72	71.51
ORF-E (lp50)	41.01	34.73	41.90	61.02	77.65		37.85	63.95	72.83	66.86	63.79	70.32
P21	31.25	49.13	47.48	26.74	36.32	32.77		61.80	59.69	55.14	58.67	58.16
Rep*-1	48.32	44.44	66.12	28.33	48.62	45.93	56.18		93.51	88.65	91.35	90.81
Rep*-2	44.60	47.70	57.28	47.37	45.52	55.49	54.59	87.57		87.44	87.62	91.58
Rep*-3	49.73	41.42	67.02	29.44	39.68	48.84	48.65	83.78	83.77		94.68	93.75
Rep*-4	50.00	47.09	60.89	37.97	52.51	44.83	53.06	85.57	83.17	42.02		94.09
Rep*-5	49.75	47.67	66.34	39.44	54.75	53.55	52.04	88.65	89.11	91.15	92.61	

[a] The upper right half of the table lists the amino acid similarity values while the lower left half lists the identity values. For each ORF-E sequence listed above the prefixes cp and lp stand for circular plasmid and linear plasmid, respectively, and the numerical designation indicates the size (in kilobase pairs) of the plasmid on which it is carried. The GenBank accession numbers for the sequences compared are as follows: *repA*, AF062395; ORF-E (cp18), U42599; ORF-E (cp32), X87127; ORF-E (lp50), X87201; *p21*, Y08413; *rep**-1, U45421; *rep**-2, U5422; *rep**-3, U45423; *rep**-4, U45424; and *rep**-5, U45425. BBF03 and BBG33 are TIGR (The Institute for Genomic Research) accession numbers.

scribed by Suk et al. [25]). The amino acid similarity values of RepA and its homologs (Table 2) ranged from 39.9 to 68.1%. All of these related proteins carry the central repeat region containing from 5 to 10 KID(E) amino acid motifs with most having between 7 and 9 KID(E) repeats (Fig. 2). As in RepA, the KID(E) motif in the RepA homologs exhibits conserved spacing of 4, 8, 15, or 22 residues. Chou-Fasman predictions indicate the majority of the repeat motif domain of

```
                                  1          2          3          4
B. turicatae 91E135 RepA     KIDSKFNELD NKIDNVENNL NTKIDTKFNE LDNKIDNVRT ELKSDIKDLD
B. burgdorferi B31 BBF03     ---------- ---------- ---------- ---------- ----------
B. burgdorferi B31 OrfE(lp50) ---------- -KIDTVEKNL NTKID----- ------NVEK NLN-------
B. burgdorferi 297 Rep-2     ---------- ---------- ---------- ----KIDSVEK NLQKDISNLD
B. burgdorferi 297 Rep-1     ---------- ---------- ---------- ---KIDSDEK NLQKDISNLD
B. burgdorferi B31 BBG33     ---------- ---------- --KIDTVKSE LTTKIDNVEK NLQKDISNLD
B. burgdorferi 297 Rep-3     ---------- ---------- --KIDNVKSE LNTKIDNVEK NLQKDISNLD
B. burgdorferi 297 Rep-4     ---------- -KIDTVEKNL NVKIDAVKSE LNTKIDNVEK NLQKDISNLD
B. burgdorferi 297 Rep-5     ---------- ---------- --KIDSVKSE LNTKIDNVEK NLNLKIDSLD
B. burgdorferi B31 OrfE(cp30) ---------- -KIDTVEKNL QKDISN---- LDIKIDAVEK NLQKDISNLD
B. burgdorferi N40 OrfE(cp18) ---------- -KIDTVEKSL NLKIDTIEKS LNLKIDFVEK SLNAKIDSLD
B. afzelii   DK1 P21         ---------- ---------- --KISDVKSE RNTKTDGLET KIDSVKNELN
                                  5          6                     7          8
B. turicatae 91E135 RepA     SKIDNVENNL NIKIDNVRTE LKSDIKDLDS KIDNVENNLN TKIDS---KF
B. burgdorferi B31 BBF03     -KIDNVKENL NLKIDNLDT- ---------- KIDTVKSELT TKIDTVKSEL
B. burgdorferi B31 OrfE(lp50) TKIDNVEKNL NLKIDNLDS- ---------- KIDTVEKNLN TKIDNVEKNL
B. burgdorferi 297 Rep-2     VKIDTVEKNL NAKIDSLDT- ---------- KIDTVEKNLN LKIDSVEKNL
B. burgdorferi 297 Rep-1     VKIDTVEKNL NAKIDSVEKN LQKDISNLDA KIDTVEKNLN AKIDSVK---
B. burgdorferi B31 BBG33     VKIDTVKSEL TTKIDNVEKN LQKDISNLDV KIDTVKSELT TKIDNVEKN-
B. burgdorferi 297 Rep-3     AKIDTVEKNL QKDISNLDI- ---------- KIDAVEKNLH VKIDTVKSEL
B. burgdorferi 297 Rep-4     TKIDNVEKNL QKDISNLDT- ---------- KIDNVEKNLQ KDI----SNL
B. burgdorferi 297 Rep-5     TKIDTVEKNL QKDISNLDI- ---------- KIDAVEKNLQ KDI----SNL
B. burgdorferi B31 OrfE(cp30) IKIDNVEKNL KKDISSLDT- ---------- KIDVVEKNLN LKIDFVEKSL
B. burgdorferi N40 OrfE(cp18) TKINNVEKTL QKDISSLDS- ---------- KIDSVKNELN SKIDSIEKTL
B. afzelii   DK1 P21         TKIDFVEKNL ETKIDGLKNE FNAGIDGLNT QIENLD---- TKID------
                                                9          10         11
B. turicatae 91E135 RepA     N---------- ---------- ---ELDNKID NVENNLNVKI E---------- --
B. burgdorferi B31 BBF03     T---------- ---------- ---------- -------TKI D---------- --
B. burgdorferi B31 OrfE(lp50) N---------- ---------- ---------- LKMDGLNIKI D---------- --
B. burgdorferi 297 Rep-2     NAKIDSLDTK IDTVEKNLQK DISNLDTKID SVEKNLNAKI D---------- --
B. burgdorferi 297 Rep-1     ---------- ---------- --SELNAKID NVEKNLNAKI D---------- --
B. burgdorferi B31 BBG33     ------LDTK IDNVEKNL-- -----DTKID NVEKNLDTKI D---------- --
B. burgdorferi 297 Rep-3     N-------TK IDNVEKNL-- -----NLKID ----GLNIKI D---------- --
B. burgdorferi 297 Rep-4     D-------TK IDNVEKNL-- -----NAKID TVEKNLNAKI D---------- --
B. burgdorferi 297 Rep-5     DI--------K IDNVEKNLQK DISNLDTKID NVEKNLNAKI DTVEKNLNTK ID
B. burgdorferi B31 OrfE(cp18) QKDISSLDNK ID-------- ---------- ---------- ---------- --
B. burgdorferi N40 OrfE(cp30) NAKIDSLDVK IDNVNNKVDY IKSELIAKID ---------- ---------- --
```

FIG. 2. Amino acid alignment of the repeat motif region of RepA and its homologs. The amino acid sequences of *Borrelia* genes exhibiting significant identity with RepA were aligned with the Pileup program and then manually adjusted to minimize gaps. The regions of these sequences that carry the repeat motifs are presented. The tripeptide KID(E) repeat motifs are highlighted by bold lettering and the individual tripeptide KID(E) repeats of *B. turicatae* 91E135 are numbered.

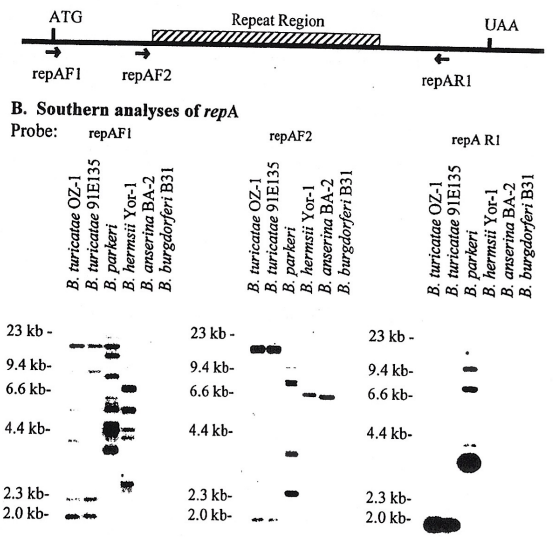

FIG. 3. Southern hybridization analyses of *repA* in different *Borrelia* species. (A) Binding sites of the *repA*-targeting oligonucleotide probes. The central repeat motif region is indicated by the boxed area. (B) Southern hybridization results obtained with *Xba*I-digested DNA and various *repA*-targeting, radioactively labeled, oligonucleotide probes (indicated above each section of panel B). Hybridization conditions were as described in the text. Molecular size standards are shown to the left of each autoradiograph.

RepA to be alpha helical. With 3.6 amino acids per turn of the helix, it is possible that all KID(E) motifs may reside on the same face of the alpha helix. The conservation of the KID(E) sequence, its repeated nature, and its conserved structural location, suggest that the repeat motif region may represent an important functional domain.

Southern blot analysis of *repA*: copy number determination. The *repA*-related genes carried by *B. burgdorferi* are either multicopy or exist in the form of gene families (23, 29). To determine if *repA* is multicopy in *B. turicatae*, Southern hybridization analyses of restricted DNA were performed with probes targeting different regions of *repA*. Through these analyses, we also sought to determine if *repA* or *repA* homologs are carried by other *Borrelia* species. To facilitate an accurate comparison of the restriction fragment length polymorphism patterns obtained with the various oligonucleotide probes that were used, multiple sets of *Xba*I-digested DNA isolated as previously described (18) were run side by side on the same 0.8% GTG-agarose gel (in standard Tris-acetate-EDTA [TAE] buffer [pH 8.0]), transferred onto Hybond N membrane (Amersham), UV cross-linked as previously described (18), and cut to generate identical Southern blots. Oligonucleotide probes (probe binding sites and sequences are indicated in Fig. 1) were 5′ end labeled by standard methods with polynucleotide kinase and [γ^{32}P]ATP (6,000 Ci/mmol; DuPont-NEN). Hybridizations were conducted at 32°C in a Hybaid hybridization oven (Labnet) by using conditions and buffers that have been previously described (20). All relapsing-fever spirochete species tested (*B. hermsii*, *B. parkeri*, and *B. turicatae*) hybridized with both the repAF1 and repAF2 probes (Fig. 3). These probes target 5′ of the central repeat motif region. The repAF1 probe hybridized with several restriction fragments in all hybridization-positive isolates. The repAF2 probe hybridized with multiple fragments in *B. turicatae* and *B. parkeri* isolates and with a single fragment in *B. hermsii* Yor-1. The detection of multiple hybridizing fragments suggests that there are several *repA*-related genes carried by the relapsing-fever spirochetes. The specificity of the hybridization of the

oligonucleotide probes is supported by the fact that many (but not all) of the hybridization-positive restriction fragments bound both *repA*-targeting probes. This observation also indicates conservation of a significant stretch of the 5′ ends of the *repA*-related genes. A probe (repAR1) targeting the 3′ end of the *repA*-related genes was also employed in hybridization analyses. This probe hybridized with multiple fragments in *B. parkeri* and one fragment in *B. turicatae* isolates. The detection of a single hybridizing fragment in the *B. turicatae* isolates with the repA-R1 oligonucleotide suggests that the 3′ ends of the *repA*-related genes may not be as conserved as the 5′ ends.

As an independent means of confirming that the multiple hybridizing fragments observed were not the result of incomplete digestion of the DNA, the blots described above were stripped and probed with an oligonucleotide (flaF1) targeting the single-copy *Borrelia fla* gene. As would be expected if complete digestion had occurred, only a single hybridizing fragment was observed for each isolate (data not shown). To further verify the hybridization results observed with the oligonucleotide probes, a PCR probe was generated from pBt2.2 with the repAF1 and repAR1 primers. With this probe, multiple hybridizing fragments were detected in several isolates. This probe also hybridized with most of the fragments that bound the repAF1 and repAF2 oligonucleotide probes (data not shown). It can be concluded from the hybridization analyses presented here that there are multiple distinct copies of *rep* in the relapsing-fever spirochetes. We refer to this group of related genes collectively as the *rep* gene family. As additional members are characterized, they can be differentiated by qualifiers such as B, C, and D, etc. It is important to note that while *B. burgdorferi* carries several *repA*-related genes, hybridization was not expected due to sequence divergence within the oligonucleotide probe target sites in the *repA* homologs of this species.

To determine if other *Borrelia* isolates and species carry *repA*-related sequences, hybridization experiments with several *repA*-targeting oligonucleotides were performed. In addition to the probes described above, one targeting the central repeat motif region was used. These probes yielded different hybridization profiles with different species (data not shown). Each of the *B. hermsii* isolates tested (YOR-1, MAN, and HS-1) hybridized with all of the *repA* probes and multiple hybridizing fragments were detected, indicating a multicopy state and general conservation of the *repA* sequence among isolates. While some of the hybridizing fragments observed among the different *B. hermsii* isolates were of the same size, others were different, indicating some divergence among isolates in and around the *repA*-related genes. In *Borrelia coriaceae*, a single hybridizing fragment was detected with the repeat motif-targeting oligonucleotide. *Borrelia anserina* hybridized strongly with the repAF2 probe, which targets just 5′ of the repeat (Fig. 3). The lack of hybridization of most of the oligonucleotide probes with DNA from these two species suggests that their *repA* genes are less conserved. While the copy number and/or composition of the *repA* gene family varies among species, it can be concluded that *repA*-related sequences are carried by numerous species of the genus *Borrelia*.

Identification of the genetic elements carrying *repA*-related sequences through Southern blot analysis of 2D CHEF-PFGE-fractionated genomic DNA. To identify the specific genetic elements that carry the multiple *repA*-related sequences, *B. turicatae* and *B. parkeri* DNA was fractionated by two-dimensional contour-clamped homogeneous electric field (Bio-Rad)-pulsed field gel electrophoresis (2D CHEF-PFGE) in 1% GTG agarose gels (15). The algorithm and parameters used were as follows: run time, 20 h 16 min; buffer, 0.5× TBE (Tris-borate-EDTA [pH 8.0]); temperature, 14°C; ramping constant, −1.400; initial switch time, 0.47 s; final switch time, 4.48 s; angle, 120°; gradient, 6 V cm^{-1}. After electrophoresis in the first dimension, the gels were rotated 90° and electrophoresed for 3 h in 0.5× TBE at 80 V (constant field). To facilitate transfer, the gels were stained with 1.0 μg of ethidium bromide ml^{-1} for 30 min, UV irradiated with 60 mJ of energy, destained, and photographed. Transfer onto Hybond-N membrane was accomplished via vacuum blotting with the VacuGene XL Vacuum Blotting System (Pharmacia) with a vacuum of 55 mbar, a 4- to 6-h transfer time, and buffers described by the manufacturer. As a consequence of the reduced mobility of circular plasmids (relative to linear plasmids) during electrophoresis, which is quite evident upon electrophoresis in the second dimension, linear and circular plasmids can be readily distinguished by 2D CHEF-PFGE (1, 5, 7, 10, 15). Upon hybridization of the blots of the 2D CHEF-PFGE-fractionated DNA, the *repA* probe hybridized with linear plasmids of 50, 35, 26, and 23 kb in both *B. turicatae* 91E135 and OZ-1 (Fig. 4). A linear plasmid of approximately 52 kb, present in 91E135 but not OZ-1, also bound the probe. In *B. parkeri*, linear plasmids of approximately 55, 50, 39, and 28 kb were hybridization positive. To verify that circular plasmids were not migrating along the axis where the linear plasmids migrate, a PCR probe targeting a *B. burgdorferi* circular plasmid-carried gene (*ospC*) was used in Southern hybridization of *B. burgdorferi* 2D CHEF-PFGE-fractionated DNA. We probed for a *B. burgdorferi* gene since circular plasmid carried genes have not been characterized from *B. turicatae*. The *ospC* probe hybridized solely with a plasmid migrating behind the axis of migration of the linear plasmids (data not shown), thereby confirming the differential migration of the plasmid conformations. These data demonstrate that *repA*-related sequences are carried on a series of linear plasmids in *B. turicatae* and *B. parkeri*. The presence of *repA* on linear plasmids is in contrast to the predominantly circular plasmid localization of the multicopy *repA* homologs (*rep*$^+$ and ORF-E) of the Lyme disease spirochetes (23, 28, 29). For example, in *B. burgdorferi* 297, six copies of *rep*$^+$ have been localized to a series of comigrating 32-kb circular plasmids (12, 23). In *B. burgdorferi* B31, copies of ORF-E are carried on a 50-kb linear plasmid and by circular plasmids of 26, 29, and 30.5 kb (28, 29). Hence, while these different bacterial species carry related genes, these *repA* homologs are carried on plasmids of different conformation.

RT-PCR analysis of *repA* expression during in vitro cultivation. To assess the transcriptional expression of *repA* in *B. turicatae* during in vitro cultivation, Northern hybridization analyses with various probes were performed (data not shown). By this approach, transcript was not detected, raising the possibility that *repA* either is not expressed during in vitro cultivation or is expressed at levels not detectable by Northern hybridization. To determine if low-level transcription was occurring, reverse transcriptase (RT) PCR was performed with an RT-PCR kit as described by the manufacturer (Perkin Elmer). The RNA template (750 ng per reaction) was purified as previously described (17) except that two rounds of RQ1-DNase digestion were performed to ensure that any contaminating DNA was digested. The repAR1 oligonucleotide served as the primer for first-strand synthesis. To synthesize double-stranded DNA, the repAF1 primer was added. The repAF1-repAR1 primer set would be expected to yield an amplicon of 666 bp and consistent with this, an amplicon of this size was obtained from *B. turicatae* OZ-1 (Fig. 5). An amplicon was not obtained from *B. turicatae* 91E135 (data not shown). Amplification products were not obtained from a minus RT, negative-control reaction. A second negative control was devised to further demonstrate that we were not amplifying residual DNA. For this control, a forward primer targeting upstream of *repA* (−303 to −284 bp upstream of the *repA* start codon) was used in conjunction with the repAR1 reverse primer in an RT-PCR reaction. Since one of these primers targets upstream of the transcriptional start site, amplification would occur only if contaminating DNA were present. Amplification was not observed, providing definitive evidence that the RNA preparations were free of contaminating DNA. To confirm that the RT-PCR amplicon from isolate OZ-1 was in fact derived from *repA*, the amplicon was cloned and partially sequenced. Partial sequence analysis of the amplicon (340 nucleotides) revealed that it contained four base differences relative to the cloned sequence from *B. turicatae* 91E135 *repA*. While it remains to be determined if these sequence differences are real or are artifacts introduced during RT-PCR, it can nonetheless be definitively concluded that the amplicon was in fact derived from amplification of a *rep* transcript.

Identification of the promoter element of *repA* by RT primer extension. To identify the putative promoter element of *repA*, RT primer extension was conducted with 5′-end-labeled OZ-PE1 primer, 750 ng of isolated RNA (17) as template, and murine leukemia virus RT (Perkin-Elmer) according to the reverse transcription protocol described above. Extension products were treated with RNase (0.5 μg μl^{-1}) (Boehringer Mannheim), extracted with phenol-chloroform-isoamyl alcohol, precipitated with ethanol, washed with 70% ethanol, vacuum dried, and resuspended in 6 μl of water. Three microliters of stop solution (Epicentre Technologies) was added to the resuspended extension products, and 3 μl was loaded onto a 6% polyacrylamide-8 M urea sequencing gel. An extension product was obtained from *B. turicatae* OZ-1 (Fig. 5) but not from 91E135. This is consistent with the RT-PCR analyses described above, which demonstrated expression of *repA* in OZ-1 but not in 91E135. From the size of the extension product, the start site could be mapped to an A residue 16 nucleotides upstream from the translational start codon (Fig. 1). Thirty nucleotides upstream from the transcriptional start site is the sequence TTG CTT, which exhibits identity with other identified *Borrelia* promoters such as those flanking *ospC* and *ospAB* (16, 17). Seventeen nucleotides downstream of the promoter is a conserved −10 or TATA box sequence element, TATACT.

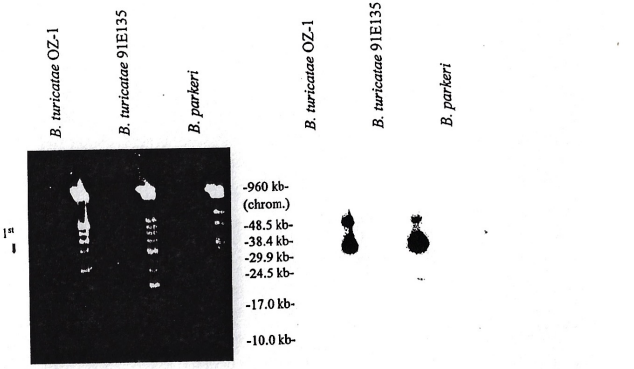

FIG. 4. Identification of the genomic elements carrying *repA* by 2D PFGE and Southern hybridization. DNA liberated from *B. turicatae* and *B. parkeri* was fractionated by 2D CHEF-PFGE as described in the text. The DNA was visualized by ethidium bromide staining and was photographed (left panel). The DNA was then transferred onto Hybond N membranes and hybridized with a PCR-generated probe as described in the text. The probe targets almost the entire *repA* gene and was generated with the repAF1 and repAR1 primers and pBT2.2, the *repA*-carrying recombinant plasmid, as amplification template. Molecular size standards (lambda monocuts) are indicated between the panels. The isolates analyzed and the direction of electrophoresis in each dimension are indicated.

Conclusions. In this report, we describe the cloning and characterization of a linear plasmid-carried gene from *B. turicatae* designated *repA*, which is a homolog of the Lyme spirochete *rep*$^+$ genes. These *rep*-related genes are characterized by the presence of a repeated potential casein kinase 2 phosphorylation site. The description "casein kinase" is broadly applied to at least two classes of ubiquitous protein kinases for which the substrates are not casein but rather a variety of enzymes and noncatalytic proteins that are involved in a variety of cellular functions. The majority of the casein kinase 2 target proteins are highly acidic (as is RepA with a pI of 4.9) and many of the phosphorylated proteins are involved in gene expression and protein synthesis (22). The conservation of sequence and spacing of the casein kinase 2 phosphorylation sites in *rep* homologs suggests that this amino acid motif may be part of an important functional domain that plays a genus-wide functional role in the biology of the *Borrelia*.

In light of what has been learned from sequence analyses of the *B. burgdorferi* genome, the presence of multiple *repA*-related sequences in the genome of other *Borrelia* species is perhaps not surprising. It is now evident that gene families comprise a significant percentage of the total number of ORFs carried by

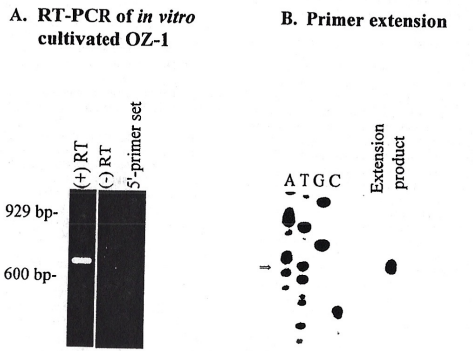

FIG. 5. Transcriptional analyses of *repA* in *B. turicatae* OZ-1 cultivated in vitro. (A) To determine if expression of *repA* occurs during cultivation in vitro, RT-PCR was performed as described in the text. In each reaction, 750 ng of isolated, DNase-treated RNA served as template. In one reaction, RT was omitted [(−)RT] to verify that all traces of DNA were removed by DNase treatment. A second negative control (5′-primer set) was also performed. In this case, a primer set targeting a region upstream of *repA* was used. (B) To identify the transcriptional start site of *repA*, primer extension analyses with the OZ-PE1 primer were performed as described in the text. The resulting primer extension products were electrophoresed in a 6% polyacrylamide-8 M urea gel alongside a sequencing ladder generated with the OZ-PE1 primer and the pBt2.2 recombinant plasmid.

the Lyme disease spirochete plasmids (1, 8, 12, 20, 23, 27). The data presented here suggest that this trend may hold true for other *Borrelia* species as well. Interestingly, the *rep*-related gene families of the Lyme disease spirochetes are present predominantly on circular plasmids, while as demonstrated in this report, in *B. turicatae* and *B. parkeri* they are present on linear plasmids. Hence, while these genes appear to be conserved, the conformation of the genetic elements that carry them is not. Similarly, while the *ospC* gene resides on a 26-kb circular plasmid in the Lyme disease spirochetes (17), in other *Borrelia* species *ospC* homologs are present on linear plasmids (19).

The conservation of the *rep* gene family and its homologous gene families among *Borrelia* species suggests that they may play an important role in the biology of the *Borrelia*. However, transcription of *repA* during in vitro cultivation could be detected only by RT-PCR and only in isolate OZ-1. These data suggest that *repA* does not play an essential role during growth in vitro. This may suggest that the functional niche of *repA* exists under other environmental conditions, perhaps during infection of mammals or in ticks. An important area of future investigation will be to assess the transcriptional activity and function of each individual *rep* allele.

The identification and characterization of proteins of unknown function that exhibit genus-wide distribution among the *Borrelia* will likely yield important information about unique aspects of *Borrelia* physiology and pathogenesis. *B. turicatae* in particular may prove to be a useful model organism for studying the functional role of *Borrelia* proteins and the factors that influence or regulate their expression. The advantage of utilizing *B. turicatae* lies in the fact that it achieves relatively high densities in the blood and tissues of infected animals. In contrast, the Lyme disease spirochetes achieve very low densities in mammals during disseminated infection and can be difficult to detect. Barbour and colleagues have recently provided a rationale and demonstration of the utility of *B. turicatae* as a model organism for the study of certain aspects of Lyme disease pathogenesis (4, 21). Through the future study of the *rep* gene family in *B. turicatae*, we hope to learn more about the potential role of related genes in other species of *Borrelia*.

Nucleotide sequence accession number. The GenBank accession number of the 786-kb ORF sequenced in this study is AF062395.

This work was supported in part by grants from the Jeffress Trust and the National Institutes of Health.

We thank the molecular pathogenesis research group at Virginia Commonwealth University for their helpful comments and suggestions. We thank Alan Barbour, Pam Pennington, and Tom Schwan for providing isolates.

REFERENCES

1. **Akins, D., S. F. Porcella, T. G. Popova, D. Shevchenko, S. I. Baker, M. Li, M. V. Norgard, and J. D. Radolf.** 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.
2. **Barbour, A. G.** 1990. Antigenic variation of a relapsing fever *Borrelia* species. Annu. Rev. Microbiol. 44:155–171.
3. **Barbour, A. G., and C. F. Garon.** 1987. Linear plasmids of the bacterium *Borrelia burgdorferi* have covalently closed ends. Science 237:409–411.
4. **Cadavid, D., D. Thomas, R. Crawley, and A. Barbour.** 1994. Variability of a bacterial surface protein and disease expression in a possible mouse model of systemic Lyme borreliosis. J. Exp. Med. 179:631–642.
5. **Carlyon, J. A., C. LaVole, S. Y. Sung, and R. T. Marconi.** 1998. Analysis of the organization of multicopy linear- and circular-plasmid-carried open reading frames in *Borrelia burgdorferi* sensu lato isolates. Infect. Immun. 66: 1149–1158.
6. **Carter, C. J., S. Bergstrom, S. J. Norris, and A. G. Barbour.** 1994. A family of surface-exposed proteins of 20 kilodaltons in the genus *Borrelia*. Infect. Immun. 62:2792–2799.
7. **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.

8. **Casjens, S., R. van Vugt, K. Tilly, P. A. Rosa, and B. Stevenson.** 1997. Homology throughout the multiple 32-kilobase circular plasmids in Lyme disease spirochetes. J. Bacteriol. 179:217–227.
9. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47:45–147.
10. **Ferdows, M. S., P. Serwer, G. A. Griess, S. J. Norris, and A. G. Barbour.** 1996. Conversion of a linear to a circular plasmid in the relapsing fever agent *Borrelia hermsii*. J. Bacteriol. 178:793–800.
11. **Fleche, A. L., D. Postic, K. Girardet, O. Peter, and G. Baranton.** 1997. Characterization of *Borrelia lusitaniae* sp. nov. by 16S ribosomal DNA sequence analysis. Int. J. Syst. Bacteriol. 47:921–925.
12. **Fraser, C., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, R. Lathigra, O. White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J. F. Tomb, R. D. Fleischman, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. Vugt, N. Palmer, M. D. Adams, J. Gocayne, J. Weidman, T. Utterback, L. Watthey, L. McDonald, P. Artiach, C. Bowman, S. Garland, C. Fujii, M. C. Cotton, K. Horst, K. Roberts, B. Hatch, H. O. Smith, and J. C. Venter.** 1997. Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*. Nature 390:580–586.
13. **Jameson, B. A., and H. Wolf.** 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. CABIOS 4:181–186.
14. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105–132.
15. **Marconi, R. T., S. Casjens, U. G. Munderloh, and D. S. Samuels.** 1996. Analysis of linear plasmid dimers in *Borrelia burgdorferi* sensu lato isolates: implications concerning the potential mechanism of linear plasmid replication. J. Bacteriol. 178:3357–3361.
16. **Marconi, R. T., M. E. Konkel, and C. F. Garon.** 1993. Variability of *osp* genes and gene products among species of Lyme disease spirochetes. Infect. Immun. 61:2611–2617.
17. **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.
18. **Marconi, R. T., D. S. Samuels, R. K. Landry, and C. F. Garon.** 1994. Analysis of the distribution and molecular heterogeneity of the *ospD* gene among the Lyme disease spirochetes: evidence for lateral gene exchange. J. Bacteriol. 176:4572–4582.
19. **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.
20. **Marconi, R. T., S. Y. Sung, C. A. N. Hughes, and J. A. Carlyon.** 1996. Molecular and evolutionary analyses of a variable series of genes in *Borrelia burgdorferi* that are related to *ospE* and *ospF*, constitute a gene family, and share a common upstream homology box. J. Bacteriol. 178:5615–5626.
21. **Pennington, P. M., C. D. Allred, C. S. West, R. Alvarez, and A. G. Barbour.** 1997. Arthritis severity and spirochete burden are determined by serotype in the *Borrelia turicatae*-mouse model of Lyme disease. Infect. Immun. 65:285–292.
22. **Pinna, L. A.** 1990. Casein kinase 2: an "eminence grise" in cellular regulation. Biochim. Biophys. Acta 1054:267–284.
23. **Porcella, S. F., T. G. Popova, D. R. Akins, M. Li, J. R. Radolf, and M. V. Norgard.** 1996. *Borrelia burgdorferi* supercoiled plasmids encode multicopy open reading frames and a lipoprotein gene family. J. Bacteriol. 178:3293–3307.
24. **Schwan, T. G., J. Piesman, W. T. Golde, M. C. Dolan, and P. A. Rosa.** 1995. Induction of an outer surface protein on *Borrelia burgdorferi* during tick feeding. Proc. Natl. Acad. Sci. USA 92:2909–2913.
25. **Suk, K., S. Das, W. Sun, B. Jwang, S. W. Barthold, R. A. Flavell, and E. Fikrig.** 1995. *Borrelia burgdorferi* genes selectively expressed in the infected host. Proc. Natl. Acad. Sci. USA 92:4269–4273.
26. **Theisen, M.** 1996. Molecular cloning and characterization of *nlpH*, encoding a novel surface-exposed, polymorphic, plasmid-encoded 33-kilodalton lipoprotein of *Borrelia afzelii*. J. Bacteriol. 178:6435–6442.
27. **Zhang, J.-R., J. M. Hardham, A. G. Barbour, and A. G. Norris.** 1997. Antigenic variation in Lyme disease Borreliae by promiscuous recombination of vmp-like sequence cassettes. Cell 89:275–285.
28. **Zuckert, W. R., R. T. Marconi, J. A. Carlyon, and J. Meyer.** 1997. Unpublished data.
29. **Zuckert, W. R., and J. Meyer.** 1996. Circular and linear plasmids of Lyme disease spirochetes have extensive homology: characterization of a repeated DNA element. J. Bacteriol. 178:2287–2298.