Species of the genus *Borrelia* cause human and animal infections, including Lyme disease, relapsing fever, and epizootic bovine abortion. The borrelial genome is unique among bacterial genomes in that it is composed of a linear chromosome and a series of linear and circular plasmids. The plasmids exhibit significant genetic redundancy and carry 175 paralogous gene families, most of unknown function. Homologous alleles on different plasmids could influence the organization and evolution of the *Borrelia* genome by serving as foci for interplasmid homologous recombination. The plasmid-carried *Borrelia direct repeat (bdr)* gene family encodes polymorphic, acidic proteins with putative phosphorylation sites and transmembrane domains. These proteins may play regulatory roles in *Borrelia*. We describe recent progress in the characterization of the *Borrelia bdr* genes and discuss the possible influence of this gene family on the biology, pathogenesis, and evolution of the *Borrelia* genome.

Species of the genus *Borrelia* cause human and animal infections, including Lyme disease, relapsing fever, and epizootic bovine abortion. The borrelial genome is unique among bacterial genomes in that it is composed of a linear chromosome and a series of linear and circular plasmids. The plasmids exhibit significant genetic redundancy and carry 175 paralogous gene families, most of unknown function. Homologous alleles on different plasmids could influence the organization and evolution of the *Borrelia* genome by serving as foci for interplasmid homologous recombination. The plasmid-carried *Borrelia direct repeat (bdr)* gene family encodes polymorphic, acidic proteins with putative phosphorylation sites and transmembrane domains. These proteins may play regulatory roles in *Borrelia*. We describe recent progress in the characterization of the *Borrelia bdr* genes and discuss the possible influence of this gene family on the biology, pathogenesis, and evolution of the *Borrelia* genome.

Species of the genus *Borrelia* cause human and animal infections, including Lyme disease, relapsing fever, and epizootic bovine abortion. The borrelial genome is unique among bacterial genomes in that it is composed of a linear chromosome and a series of linear and circular plasmids. The plasmids exhibit significant genetic redundancy and carry 175 paralogous gene families, most of unknown function. Homologous alleles on different plasmids could influence the organization and evolution of the *Borrelia* genome by serving as foci for interplasmid homologous recombination. The plasmid-carried *Borrelia direct repeat (bdr)* gene family encodes polymorphic, acidic proteins with putative phosphorylation sites and transmembrane domains. These proteins may play regulatory roles in *Borrelia*. We describe recent progress in the characterization of the *Borrelia bdr* genes and discuss the possible influence of this gene family on the biology, pathogenesis, and evolution of the *Borrelia* genome.

Address for correspondence: Richard T. Marconi, Department of Microbiology and Immunology, Medical College of Virginia at VCU, Richmond, VA 23298-0678, USA; fax: 804-828-9946; e-mail: Rmarconi@hsc.vcu.edu.

The *bdr* Gene Families of the Lyme Disease and Relapsing Fever Spirochetes: Potential Influence on Biology, Pathogenesis, and Evolution

David M. Roberts,* Jason A Carlyon,† Michael Theisen,‡ and Richard T. Marconi*

*Medical College of Virginia at Virginia Commonwealth University, School of Medicine, Richmond, Virginia, USA; †Yale School of Medicine, Yale University, New Haven, Connecticut, USA; and ‡Statens Serum Institute, Copenhagen, Denmark

Emerging Infectious Diseases Vol. 6, No. 2, March–April 2000

110
and is rare in the United States. Endemic relapsing fever is more prevalent, predominantly in the western regions. Three closely related *Borrelia* species, *B. hermsii*, *B. turicatae*, and *B. parkeri*, are associated with this disease. Hallmark features of relapsing fever include cyclic fever and spirochtemia. The molecular basis for these features can be attributed to the differential production of dominant variable surface antigens of the Vmp protein families (16). The 40 or so plasmid-carried *vmp* related genes in the *B. hermsii* genome are expressed only one at a time. A single expression locus exists, and genes not at this site lack a promoter element and are therefore not transcribed (17). The expressed Vmp becomes a primary target of a vigorous humoral immune response that kills most of the spirochetal population. However, at a frequency of approximately of $1 \times 10^{-3}$ to $1 \times 10^{-4}$ per generation, the identity of the expressed Vmp changes (18) through gene conversion (19). The net effect of this nonreciprocal event is to replace the gene located in the expression locus with one that was previously silent. Production of a new antigenically distinct Vmp allows evasion of the humoral immune response. This ongoing change in Vmp synthesis allows the relapsing fever spirochete population to reestablish itself in the host, thus leading to spirochetemia and the relapse of fever. Antigenic variation systems have also been identified in the Lyme disease spirochetes; however, they appear to exert a more subtle effect (20).

While clinical relapsing fever and Lyme disease differ from each other in many ways, their causative agents share many similarities at both the biologic and genetic levels. At the biologic level, they are host associated and undergo similar environmental transitions in the course of cycling between mammals and arthropods. In view of the distinctly different characteristics of these environments, the spirochetes must be able to adapt rapidly. Evidence suggests that the relapsing fever and Lyme disease spirochetes use related proteins to adapt to or carry out similar functions in changing environments. For example, homologs of the plasmid-carried *ospC* gene of the Lyme disease spirochetes are carried by several other *Borrelia* species, including the relapsing fever spirochetes (21). Both *ospC* and its relapsing fever spirochete homolog (*vmp33*) are selectively expressed during the early stages of infection, which suggests that they play a common functional role (22,23). The *B. burgdorferi* Rep or Bdr protein family is also distributed genuswide. Members of this polymorphic protein family possess highly conserved putative functional motifs and structural properties, which suggests that they may also carry out an important genuswide role (24,25).

### The *Borrelia* Genome

At the molecular level, a unique feature of *Borrelia* is the unusual organization and structure of their genome. Unlike most bacteria, which carry their genetic material in the form of a single, circular DNA molecule, *Borrelia* have a segmented genome (26-28). Most genetic elements carried by these bacteria are linear with covalently closed termini or telomeres (27). The telomeres are characterized by short hairpin loops of DNA (29). If heat denatured, these linear molecules relax to form a single-stranded circular molecule. If reannealed, they base-pair upon themselves to form a double-stranded linear molecule that by physical necessity possesses a short single-stranded hairpin loop at each telomere. Genetic elements of this structure are rare in bacteria and are reminiscent of certain viral genomes. In *B. burgdorferi* (isolate B31), the largest of the linear genomic elements is the 911-kb chromosome (30). The chromosome carries 853 putative ORFs, most of which are thought to encode housekeeping functions. The remaining 12 linear and 8 circular genetic elements are plasmids. The plasmids might best be thought of as mini-chromosomes, since as a group they are indispensable in situ and may carry genes encoding proteins involved in housekeeping functions (31). In addition, they may further deviate from the true definition of a plasmid in that their replication may not be independent and may instead be tightly coordinated with the replication of the chromosome (32,33).

Nearly 50% of the plasmid-carried ORFs lack homology with known sequences, which suggests that their encoded proteins may define the unique biologic and pathogenetic aspects of *Borrelia* (30). Several of the proteins derived from these plasmid-carried genes of unknown function are antigenic or selectively expressed during infection, which indicates that they function in the mammalian environment (20,34-37). A striking feature of the plasmid-carried ORFs is that they are organized into 175 paralogous gene families of two or more members (30). Hence, the
Synopses

DNA content of the plasmids is highly redundant. Since the maintenance of DNA is energetically expensive, it is likely that this redundant DNA is of biologic importance to Borrelia. The paralogous gene families of Borrelia have been the focus of intensive research as they are thought to play important roles in pathogenesis and to influence genome organization and evolution (20,30,35,38-40).

Table 1. Borrelia species carrying bdr-related genes or expressing proteins immunoreactive with anti-Bdr antisera

<table>
<thead>
<tr>
<th>Species</th>
<th>Associated disease</th>
<th>Arthropod vector</th>
<th>Bdr-related information</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. burgdorferi</td>
<td>Lyme disease, endemic worldwide</td>
<td>I. scapularis, I. ricinus, I. pacificus</td>
<td>All bdr-gene family members (18 total) have been identified in isolate B31G (30), bdr-alleles that are organized into 3 subfamilies (D,E,F) (25), the genes are carried on variably sized linear and circular plasmids (30)</td>
</tr>
<tr>
<td>B. afzelii</td>
<td>Lyme disease, Eurasia</td>
<td>I. ricinus, I. persulcatus</td>
<td>Single bdr gene has been sequenced (43); several Bdr-related proteins have been detected by immunoblot analysis (this report)</td>
</tr>
<tr>
<td>B. garinii</td>
<td>Lyme disease, Eurasia</td>
<td>I. ricinus, I. persulcatus</td>
<td>Bdr proteins detected by immunoblot analyses only (this report)</td>
</tr>
<tr>
<td>B. tanukii</td>
<td>Not associated with human disease</td>
<td>I. pacificus, I. scapularis, I. spinipalpis</td>
<td>Bdr proteins detected by immunoblot analyses only (this report)</td>
</tr>
<tr>
<td>B. bissettii</td>
<td>Not associated with human disease</td>
<td>I. pacificus</td>
<td>Bdr proteins detected by immunoblot analyses only (this report)</td>
</tr>
<tr>
<td>B. andersonii</td>
<td>Not associated with human disease</td>
<td>I. dentatus</td>
<td>Bdr genes detected by hybridization and Bdr-related proteins by immunoblot analyses (this report, data not shown)</td>
</tr>
<tr>
<td>B. valaisiana</td>
<td>Not associated with human disease</td>
<td>I. columnae, I. ricinus, I. ovatus</td>
<td>Bdr proteins detected by immunoblot analyses only (this report)</td>
</tr>
<tr>
<td>B. japonica</td>
<td>Not associated with human disease</td>
<td>I. ovatus</td>
<td>Bdr proteins detected by immunoblot analyses only (this report)</td>
</tr>
<tr>
<td>B. hermsii</td>
<td>Endemic relapsing fever, United States</td>
<td>Onithodoros hermsii</td>
<td>Numerous bdr genes have been described and are carried on both linear and circular plasmids (25,46); several Bdr proteins have been detected by immunoblot analyses (this report) (44)</td>
</tr>
<tr>
<td>B. parkeri</td>
<td>Endemic relapsing fever, United States</td>
<td>Onithodoros parkeri</td>
<td>Two bdr-related genes have been sequenced (25), and others have been detected by hybridization with genes residing on both linear and circular plasmids (46); several Bdr proteins detected by immunoblot analyses (this report) (44)</td>
</tr>
<tr>
<td>B. turicatae</td>
<td>Endemic relapsing fever, United States</td>
<td>O. turicata</td>
<td>At least nine bdr-related genes have been described and are present on linear plasmids ranging from 25 to 220 kb in size (24,46); several Bdr proteins have been detected by immunoblotting (this report) (24,44)</td>
</tr>
<tr>
<td>B. miyamotoi</td>
<td>Relapsing fever?</td>
<td>I. persulcatus</td>
<td>Bdr proteins detected by immunoblotting only (this report)</td>
</tr>
<tr>
<td>B. coriaceae</td>
<td>Epizootic bovine abortion, United States</td>
<td>O. coriaceus</td>
<td>bdr-related genes and proteins detected by hybridization (46) or immunoblotting (this report)</td>
</tr>
<tr>
<td>B. anserina</td>
<td>Avian spirochetosis, United States</td>
<td>Argas persicus</td>
<td>bdr-related sequences have been detected by hybridization (46); Bdr-related proteins were not detected in in vitro cultivated bacteria (this report)</td>
</tr>
</tbody>
</table>

Identification of Borrelia Direct Repeat (bdr) Related Genes

The bdr gene family is a large, polymorphic, plasmid-carried, paralogous gene family of unknown function that was originally identified in B. burgdorferi (41,42). Members of this gene family have been characterized in several Borrelia species and isolates (Table 1) and have
been assigned various gene names (25,41-44) (Table 2). We have adopted the \( bdr \) designation in the context of a nomenclature system (25), summarized below. Genes belonging to the \( bdr \) gene family were first identified through the analysis of repeated DNA sequences in \( B. burgdorferi \) sensu lato complex isolates (41,42). Seven nonidentical but closely related copies of a plasmid-carried repeated element were identified in \( B. burgdorferi \) 297 (42). Three additional copies of this repeated sequence were further identified in \( B. burgdorferi \) 297 (45). These loci carry several ORFs that were designated as \( rep^+ \), \( rep^- \), LPA, LPB (the LP genes have recently been redesignated as \( mlp \) for multicopy lipoprotein [45]), \( rev \), and the \( orf4BCD \) operon (note: ORFs A and B have been redesignated as \( blyA \) and \( blyB \)). Some of these genes, particularly \( rep \) and \( mlp \), exhibit allelic variation and encode polymorphic proteins, the functions of which are under investigation. Focusing specifically on the \( rep \) or \( bdr \) genes, the \( rep \) designation was originally

### Table 2. *Borrelia* Bdr homology groups and gene nomenclature

<table>
<thead>
<tr>
<th>Bdr subfamily</th>
<th>Species/revised gene designation</th>
<th>Accession or TIGR number</th>
<th>Previous gene names</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subfamily A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. turicatae OZ-1</td>
<td>( bdrA )</td>
<td>AF062395</td>
<td>repA</td>
<td>(46)</td>
</tr>
<tr>
<td>B. turicatae OZ-1</td>
<td>( bdrA_1,A_2,A_3 )</td>
<td>AF128445-AF128447</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. hermsii YOR-1</td>
<td>( bdrA_1,A_2,A_3 )</td>
<td>AF143473-AF143475</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. hermsii HS1</td>
<td>( bdrA_1,A_2 )</td>
<td>AF143457-AF143458</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. hermsii MAN</td>
<td>( bdrA_1,A_2 )</td>
<td>AF143465, AF143467</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. parkeri</td>
<td>( bdrA_1 )</td>
<td>AF143455</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>Subfamily B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. turicatae OZ-1</td>
<td>( bdrB_1,B_2,B_3,B_4,B_5 )</td>
<td>AF128448-AF128452</td>
<td>none</td>
<td>(24)</td>
</tr>
<tr>
<td>B. hermsii MAN</td>
<td>( bdrB_1,B_2,B_3 )</td>
<td>AF143463, AF143464, AF143466</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>Subfamily C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. parkeri</td>
<td>( bdrC_1 )</td>
<td>AF143455</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. hermsii MAN</td>
<td>( bdrC_1,C_2,C_3,C_4,C_5 )</td>
<td>AF143468-AF143472</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. hermsii HS1</td>
<td>( bdrC_1,C_2,C_3,C_4 )</td>
<td>AF143459-AF143462</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. hermsii YOR-1</td>
<td>( bdrC_1 )</td>
<td>AF143476</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>B. parkeri</td>
<td>( bdrC_2 )</td>
<td>AF143456</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>Subfamily D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. burgdorferi B31G</td>
<td>( bdrD_1,D_2,D_3 )</td>
<td>BBL35, BBM34, BBO34</td>
<td>( bdrO, bdrK, bdrM )</td>
<td>(30)</td>
</tr>
<tr>
<td>B. burgdorferi B31G</td>
<td>( bdrD_2,D_5,D_6 )</td>
<td>BBP34, BBQ42, BBS37</td>
<td>( bdrA, bdrV, bdrE )</td>
<td>(30)</td>
</tr>
<tr>
<td>B. burgdorferi B31</td>
<td>( bdrD_2 )</td>
<td>X87201</td>
<td>ORF-E (lp50 allele)</td>
<td>(41)</td>
</tr>
<tr>
<td>B. burgdorferi B31</td>
<td>( bdrD_3 )</td>
<td>X87127</td>
<td>ORF-E (cp30.5 allele)</td>
<td>(41)</td>
</tr>
<tr>
<td>B. burgdorferi B31</td>
<td>( bdrD_4 )</td>
<td>U42599</td>
<td>ORF-E (cp18 allele)</td>
<td>(41)</td>
</tr>
<tr>
<td>B. burgdorferi B31</td>
<td>( bdrD_6 )</td>
<td>BBN34</td>
<td>( bdrQ )</td>
<td>(30)</td>
</tr>
<tr>
<td>B. burgdorferi B31</td>
<td>( bdrD_{10} )</td>
<td>BBR35</td>
<td>( bdrG )</td>
<td>(30)</td>
</tr>
<tr>
<td>Subfamily E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. burgdorferi B31G</td>
<td>( bdrE_1,E_2,E_3 )</td>
<td>BBL27, BBN27, BBO27</td>
<td>( bdrP, ) none, ( bdrV )</td>
<td>(30)</td>
</tr>
<tr>
<td>B. burgdorferi B31G</td>
<td>( bdrE_2,E_6 )</td>
<td>BBR27, BBS29, BBQ34</td>
<td>( bdrH, ) ( bdrF, ) ( bdrW )</td>
<td>(30)</td>
</tr>
<tr>
<td>B. burgdorferi 297</td>
<td>( bdrE_1,E_2 )</td>
<td>U45421, U45422</td>
<td>( rep+2.9-1, rep+2.9-2 )</td>
<td>(42)</td>
</tr>
<tr>
<td>B. burgdorferi 297</td>
<td>( bdrE_4 )</td>
<td>U45423, U45424</td>
<td>( rep+2.9-3, rep+2.9-4 )</td>
<td>(42)</td>
</tr>
<tr>
<td>B. burgdorferi 297</td>
<td>( bdrE_5 )</td>
<td>U45425</td>
<td>( rep+2.9-5 )</td>
<td>(42)</td>
</tr>
<tr>
<td>B. burgdorferi 297</td>
<td>( bdrE_6 )</td>
<td>AF046998</td>
<td>( rep+2.9-8 )</td>
<td>(45)</td>
</tr>
<tr>
<td>B. burgdorferi 297</td>
<td>( bdrE_7 )</td>
<td>AF046999</td>
<td>( rep+2.9-9 )</td>
<td>(45)</td>
</tr>
<tr>
<td>Subfamily F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. afzelii DK1</td>
<td>( bdrF_1 )</td>
<td>Y08143</td>
<td>p21</td>
<td>(43)</td>
</tr>
<tr>
<td>B. burgdorferi B31G</td>
<td>( bdrF_1,F_2,F_3 )</td>
<td>BBF03, BBG33, BBH13</td>
<td>( bdrS, ) ( bdrT, ) ( bdrU )</td>
<td>(30)</td>
</tr>
</tbody>
</table>
chosen to reflect a central repeat motif carrying domains in the deduced amino acid sequences. The + and - designations were assigned to indicate that the overlapping rep+ and rep- genes are located on opposing DNA strands. Plasmid-carried repeated DNA sequences were also identified in B. burgdorferi B31 and found to carry either all or a subset of seven ORFs, designated A through G (41). Of relevance to this discussion are the ORF-E sequences that are rep or bdr homologs. A bdr-related gene was also identified in B. afzelii DK1 and designated as p21 (43). B. afzelii causes Lyme disease in Europe and Asia. The rep+, ORF-E, and p21 designations have recently been replaced with bdr gene designations (24,25,44).

To assess and compare the composition and complexity of the bdr gene family among species and isolates of the B. burgdorferi sensu lato complex, restriction fragment length polymorphism (RFLP) patterns were determined (Appendix). Genomic DNA digested with Xba1 was Southern blotted and probed with an oligonucleotide targeting the bdr genes (Figure 1). A variable number of hybridizing bands of different size were detected. These analyses demonstrate that extensive bdr gene families are carried by B. burgdorferi sensu lato complex isolates and that the RFLP patterns vary at the inter- and intraspecies level. Hybridization analyses of other Borrelia species showed that they also carry bdr-related gene families (24,25,46). bdr-related genes have been detected by hybridization in B. turicatae, B. hermsii, B. parkeri, B. coriaceae, and B. anserina (25,46). Isolates of these species also exhibit substantial variation in their bdr RFLP patterns at the intraspecies level. Table 1 lists the Borrelia species that carry bdr-related genes and indicates the methods by which these genes or proteins were detected.

Sequences flanking some bdr alleles also appear to be distributed genus wide. Some bdr alleles of B. turicatae, B. parkeri, and B. hermsii are flanked by genes that are homologs of genes carried by the Lyme disease spirochetes (24,25). As a specific example, the B. turicatae bdrA₁ gene is flanked by ORFs that are homologs of the BBG34 and BBG30 genes of B. burgdorferi (24,25). In the Lyme disease spirochetes, BBG34 is part of a three-member paralogous gene family, while BBG30 is a single-copy gene (30). Located between BBG30 and BBG34 is BBG33, a member of the bdr gene family (recently redesignated as bdrF₂) (25). Although these divergent Borrelia species carry related genes, their organization differs (24), which indicates that rearrangement has taken place in the ancestral plasmid that carried these homologs. Figure 2 compares the
Evolutionary Analyses of \textit{bdr}-Related Sequences: Revised Nomenclature for the Bdr-Related Proteins

To simplify the complicated nomenclature of \textit{bdr}-related genes, a \textit{bdr} nomenclature system has been developed that assigns gene names on the basis of phylogenetic relationships inferred from comparative analysis of genetically stable regions of the \textit{bdr} genes (25). This system, which is applicable genuswide, allows for a ready assessment of relationships among \textit{bdr} paralogs and orthologs. The rationale for this system stemmed from the results of a comprehensive evolutionary analysis of $>50$ \textit{bdr}-related sequences from five \textit{Borrelia} species that demonstrated that \textit{bdr} sequences are organized into six distinct subfamilies, designated A through F (25). Subfamilies are not necessarily species specific; some contain \textit{bdr} alleles from different \textit{Borrelia} species (25). Since members of a given subfamily are closely related to one another with identity values for the N terminal domain being $>95\%$, each member is assigned the same gene name designation, and paralogs are distinguished by a numerical subscript. In \textit{B. turicatae} OZ-1, two \textit{bdr} subfamilies, \textit{bdrA} and \textit{bdrB}, contain at least four and five members, respectively (24). Members of the \textit{bdrA} subfamily are designated \textit{bdrA}$_1$, \textit{bdrA}$_2$, \textit{bdrA}$_3$, and \textit{bdrA}$_4$ while members of the \textit{bdrB} family are designated \textit{bdrB}$_1$ through \textit{bdrB}$_5$. This revised Bdr nomenclature scheme was modeled after that proposed for bacterial polysaccharide synthesis genes (47) and is in accordance with the nomenclature guidelines established by Demerec (48).

The subfamily affiliation of \textit{bdr} genes can be readily determined through comparative sequence analyses of the amino acid segment preceding the polymorphic repeat motif region of these proteins (described in detail below) (25). Relationship assessments based on the genetically stable N terminal domain (vs. complete sequences) are preferable because the calculated evolutionary distances and clustering relationships are not artificially skewed by the variable number of repeat motifs present in the repeat motif domain. Since the genetically unstable repeat motif domain comprises as much as 50\% of the total coding sequence in some alleles, it can
have a substantial impact on inferred relationships. In addition, extensive sequence variation in the carboxyl termini of the Bdr proteins at the inter-species level makes it difficult to align this domain with confidence, which further influences the inferred relationships.

*bdr* evolutionary analyses show that *Borrelia* species carry members of at least two *bdr* subfamilies (25,44). In fact, *B. burgdorferi* carries three distinct subfamilies. Multiple Bdr subfamilies in diverse *Borrelia* species suggest that there has been selective pressure to maintain multiple *bdr* alleles and *bdr* genetic diversity. This genetic diversity may increase the functional diversity of the Bdr proteins.

**Molecular Features and Physical Properties of the Bdr Proteins**

While early analyses of *Borrelia bdr* genes demonstrated their multicopy nature (41,42,46), the full extent of the complexity of the *bdr* gene family in the Lyme disease spirochetes was not fully recognized until the *B. burgdorferi* genome sequence was determined (30). *B. burgdorferi* B31 was found to carry 17 distinct *bdr*-related genes (and one truncated variant) distributed among different linear and circular plasmids. *B. turicatae*, which carries at least nine different *bdr* alleles, carries these genes exclusively on linear plasmids (24,25,46). Other relapsing fever spirochete species (*B. parkeri* and *B. hermsii*) are similar to the Lyme disease bacteria in that they carry *bdr* genes on both linear and circular plasmids (25). In the Lyme disease spirochetes each of the 32-kb circular plasmids, with the exception of plasmids M and P, carry two different *bdr* genes separated by seven or eight ORFs. Each of these circular plasmids carries one *bdrD* subfamily member and one *bdrE* subfamily member. The maintenance of genes belonging to different subfamilies on a single plasmid is consistent with the possibility that each carries out a different function. In contrast, in the Lyme disease spirochetes, the *bdrF* subfamily members are localized to linear plasmids with only a single *bdr* gene per plasmid. These observations suggest that there has been selective pressure to maintain the association of specific subfamilies with specific types of plasmids. Less is known about the *bdr*-carrying plasmids and the organization of the *bdr* genes and subfamilies in the relapsing fever borreliae. However, as in the Lyme disease spirochetes, in *B. turicatae* most *bdr*-carrying plasmids carry two *bdr* genes, one from subfamily *bdrA* and one from subfamily *bdrB* (24).

The sequence of more than 50 *bdr* alleles from five different *Borrelia* species has been determined (Table 2) (24,25,41-43,46). These extensive comparative sequence analyses led to the identification of conserved features that provide insight into the possible biologic roles of the Bdr proteins. For example, all *bdr* alleles carry centrally located repeat motif domains (Figure 3). Although conserved in sequence, these domains vary in length among alleles as a result of varying numbers of the repeat motif. The core tripeptide

---

Figure 3. Key features and putative functional domains of the Bdr proteins. The schematic depicts a prototype Bdr protein with the characteristics of each domain indicated. The abbreviation, ID%, is for percentage amino acid identity at either the inter- or intra-family level as indicated in the figure. Standard amino acid abbreviations are used in the figure to denote the conserved C-terminal lysine (K) or asparagine (N) residues, which are thought to be exposed in the periplasm and the cytoplasmically located core tripeptide of the repeat motif (lysine-isoleucine-aspartic acid; KID).
the repeat is the sequence KID. The repeat motifs encode consensus casein kinase 2 phosphorylation (CK2P) motifs of the sequence T/SKID/E (43). While it may appear somewhat paradoxical for bacteria to carry casein kinases, casein kinase is a descriptive term broadly applied to at least two classes of ubiquitous protein kinases for which the substrates may include various enzymes and noncatalytic proteins involved in important cellular regulatory functions (49). Most proteins phosphorylated by CK2-like kinases are highly acidic, as are the *Borrelia* Bdr proteins (isoelectric points between 5 and 6). The phosphorylation site in CK2P motifs is either the Ser or Thr residue of the motif. Although histidine kinases have been known to exist in some bacteria, it has been widely held that bacteria lack Ser - Thr kinases. However, Ser - Thr kinases have recently been identified in several bacterial species, including *Myxococcus, Anaebana, Freymella, Yersinia,* and *Streptomyces* (50). Most importantly, analysis of the *B. burgdorferi* genome sequence identified a putative Ser - Thr kinase designated BB0648 (30,50). This ORF carries a domain that exhibits homology with the active site of Ser - Thr kinases. *B. burgdorferi* also carries a homolog of the PPM family of eucaryotic protein Ser - Thr phosphatases (30,50). The presence of these genes in *B. burgdorferi* suggests that the *Borrelia* possess the machinery necessary for Ser - Thr phosphorylation and dephosphorylation.

Another important conserved feature identified through sequence analyses is the hydrophobic carboxyl terminal domain of approximately 20 amino acids. Computer analyses conducted with the TMpred program indicate that this domain has a high propensity to form a transmembrane helix (24,25). The Tmpred values for the 20 aa C-terminal domains are 2,000 to 2,600. A value of 500 or greater is considered significant (24,25). Comparison of the Bdr putative transmembrane domain sequences from the Lyme disease spirochetes with those from the relapsing fever spirochetes indicates that, while there is conservation in physical properties, there is essentially no conservation of primary sequence. However, sequence conservation does exist at the subfamily level (24,25). Since the Bdr proteins lack an obvious export signal, membrane association would most likely be with the spirochetal inner membrane, with the rest of the protein, which is hydrophilic, extending into the cytoplasm. The terminal residue of the protein is in almost all cases a positively charged amino acid (lysine or asparagine). This residue could extend into the periplasm and serve to anchor the Bdr proteins to other cellular components, such as the peptidoglycan.

**Immunologic Analyses of the Bdr Proteins**

The presence of multiple *bdr* alleles and *bdr* subfamilies within isogenic populations has prompted speculation that there may be differential expression at either the subfamily or individual allele level, possibly in response to environmental stimuli (46). Limited studies of *bdr* expression and production, based on either mRNA detection or immunoblot analyses, have been performed. Porcella et al. (42) used Northern hybridization to determine if expression of *B. burgdorferi* *bdr*-related genes occurs during cultivation in the laboratory under standard culture conditions (33°C in BSK media). *Bdr* transcripts were not detected by this approach. Similarly, in an earlier analysis, we also conducted Northern hybridization experiments to assess *bdr* expression (46). We detected expression of *B. turicatae* OZ1 *bdrA* subfamily members in bacteria cultivated under standard laboratory growth conditions (46). However, when reverse transcriptase (RT)-PCR methods were applied, transcription of a single *bdrA* allele was detected (46). *B. turicatae* OZ-1 was later demonstrated to carry at least nine *bdr* alleles, four of which belong to the *bdrA* subfamily. Analysis of the sequence of these alleles showed that all four should have been readily amplified by the RT-PCR primer set because of the conservation of the primer binding sites (24). The lack of detection of transcript derived from these alleles suggested that only a subset of the *bdr* A subfamily alleles is expressed. This raised the possibility that other *bdr* alleles are either nonfunctional genes or their expression requires different environmental stimuli. The transcriptional expression of the *bdrB* subfamily has not been specifically assessed. Thorough transcriptional analyses using allele-specific probes and primers are an important step, since they allow specific assessment of the expression of individual *bdr* alleles under differing environmental conditions. In addition, analyses of the upstream DNA sequences of individual *bdr* alleles and their genomic location may elucidate the molecular basis for *bdr* transcriptional regulation.
Immunologic analyses have provided a somewhat different overall picture regarding Bdr production. Immunologic analyses described in this report and elsewhere (44) demonstrate that several members of the bdr gene family are expressed during in vitro cultivation. We conducted a comprehensive analysis of the expression of Bdr proteins among *Borrelia* species. When antisera raised against recombinant *B. afzelii* BdrF1 (24) were used in immunoblot analyses, several immunoreactive proteins were detected in cell lysates of all *Borrelia* species tested (Figure 4). The only exception was *B. anserina*, a causative agent of avian spirochetosis. Although *bdr*-related sequences have been detected in *B. anserina* by hybridization techniques (46), immunoreactive proteins were not detected in immunoblot analyses. Additional analyses are required to determine if this indicates absence of translational expression or the lack of epitope conservation in this species. In any event, the fact that immunoreactive bands were not detected in this species attests to the specificity of the anti-Bdr antisera. As a further demonstration of the specificity of the antisera and to highlight the fact that the Bdr proteins are unique to *Borrelia*, a cell lysate of *Leptospira interrogans* was included in the immunoblot analyses. Immunoreactivity with proteins in the Bdr size range was not observed with the anti-Bdr antisera in this spirochete species. *Borrelia* species that expressed immunoreactive proteins included *B. garinii*, *B. burgdorferi*, *B. turdae*, *B. tanukii*, *B. japonica*, *B. valaisiana*, *B. afzelii*, *B. coriaceae*, *B. bissetii*, *B. miyamotoi*, *B. parkeri*, *B. hermsii*, and *B. turicatae* (Table 1). Particularly striking was the extensive variation in the number and molecular weight of the immunoreactive proteins expressed, with up to 12 distinct Bdr proteins detected. Variation in expression patterns was observed at both the inter- and intraspecies level. Analysis of three *B. burgdorferi* isolates (B31G, cN40, and CA12) demonstrated variability in both the size and number of expressed Bdr proteins. Isolate B31G has been demonstrated by genomic sequencing to carry 18 distinct bdr alleles. Immunoblot analyses show that not all alleles are expressed during in vitro cultivation; therefore, some alleles may be differentially regulated.

Figure 4. Immunoblot analyses demonstrating the variation in Bdr protein expression in *Borrelia* species and isolates. Bacteria were cultivated and prepared for analysis as described in the methods. Proteins were fractionated by SDS-PAGE, immunoblotted and screened with anti-BdrF1-*B. afzelii* DK1 antisera. The species and isolates analyzed are indicated above each lane in panels A, B and C. The migration positions of the protein standards are indicated in each panel.
The broad immunoreactivity of the antisera with diverse *Borrelia* species indicates that some epitopes are conserved genuswide. In view of the sequence divergence in the N and C terminal domains of the Bdr proteins derived from different subfamilies, it is likely that the cross-reactive epitopes reside in the conserved repeat motif region. Consistent with this, computer analyses of the repeat domain of all determined Bdr protein sequences predict them to be alpha helical and to have a surface exposed on the protein and a positive Jameson-Wolf antigenic index (24,25,44). The conservation and synthesis of these polymorphic proteins in such a diverse group of *Borrelia* species suggest that they play an important role in *Borrelia* biology genuswide.

**The Bdr Proteins and *Borrelia* Biology: An Overview**

*Bdr* genes and extensive *bdr* gene families have now been identified and characterized in several diverse *Borrelia* species (24,25,42-44,46). Comparative sequence analyses, which have identified conserved putative functional domains, have provided the basis for the development of hypotheses regarding Bdr function and cellular location. The Bdr proteins, which lack known consensus export signals, are likely anchored to the cytoplasmic membrane through their conserved, hydrophobic, putative transmembrane spanning domain. The C-terminal positively charged amino acid may be exposed to the periplasm, where it may interact with other cellular components that may include the peptidoglycan. The repeat motif domain, which is predicted by computer analyses to be hydrophilic and surface exposed on the protein, likely extends into the cytoplasm. The conserved repeat motif domain that carries the putative Ser-Thr phosphorylation motifs may then be accessible for phosphorylation or to interact with other cytoplasmic proteins or DNA to form a membrane anchored complex. As with numerous other proteins, phosphorylation and dephosphorylation could play a regulatory role, perhaps in signaling or sensing.

Multiple polymorphic *bdr* alleles may also allow or lead to the continual modification of these genes through homologous recombination. In fact, the variable nature of the repeat motif region, which is clearly not evolutionarily stable, has likely arisen from slipped-strand mispairing, recombination, or rearrangement. In view of the extensive genetic redundancy of the plasmid component of the *Borrelia* genome, recombination in and among related sequences on different plasmids could affect the organization and evolution of the genome and ultimately host-pathogen interaction. Inter- or intra-plasmid exchange of DNA sequences could provide a mechanistic basis for the extensive genetic variability that has been widely described for *Borrelia* plasmids (28,29,51-59). In spite of the apparent necessity for at least most of the plasmids for survival, as inferred from their ubiquitous distribution among *Borrelia* isolates, these bacteria are able to tolerate remarkable genomic variability. Diversity in the plasmids and the genes they carry may actually be exploited as a tool for phenotypic diversity and rapid environmental adaptation.

**Acknowledgments**

We thank Michael Norgard, Steve Porcella, Justin Radolf, and the Molecular Pathogenesis research group at Virginia Commonwealth University for helpful discussions. This work was supported in part by a grant from the Jeffress Memorial Trust.

Mr. Roberts is a 3rd-year Ph.D. student, Department of Microbiology and Immunology, Medical College of Virginia at Virginia Commonwealth University. His scientific interests are centered on the study of the molecular mechanisms of pathogenesis and adaptive responses of arthropod-carried spirochetes, specifically on the role of plasmid-carried gene families in *Borrelia* biology and pathogenesis.

**References**

Synopses


Appendix

Bacterial Cultivation, DNA Isolation, and Southern Hybridization Analyses

Isolates belonging to the *Borrelia burgdorferi* sensu lato complex were cultivated in complete BSK-H media (Sigma) at 33°C. To cultivate the relapsing fever borreliae and other *Borrelia* species, the complete BSK-H media were supplemented with additional rabbit sera (Sigma) to a final concentration of 12% (vol/vol). Bacteria were harvested by centrifugation and washed with phosphate buffered saline (pH 7.0), and DNA was extracted (25). For Southern hybridization analyses, 5 g of DNA from each isolate was digested under standard conditions with *XbaI* and fractionated by electrophoresis in 0.8% GTG agarose gels. (The DNA was transferred onto membranes for hybridization by vacuum blotting using the VacuGene system as described by the manufacturer (Pharmacia). All other Southern hybridization methods were as previously described (39).
**Immunoblot Analyses**

Bacterial cultures were grown and harvested as described above. One OD600 equivalent of cells was pelleted and resuspended in 100 l of standard SDS-sample buffer with reducing agents. The cell lysates (7 l) were fractionated by electrophoresis in 15% SDS-PAGE gels and electroblotted onto Immobilon P membranes (38). The immunoblots were blocked overnight in blocking buffer (1X PBS, 0.2% Tween, 0.002% NaCl, and 5% nonfat dry milk) and then incubated with a 1:1,000 antisera dilutions. ImmunoPure Goat anti-mouse IgG (H+L) peroxidase conjugate served as the secondary antibody. The secondary antibody was incubated with the blots for 1 hour at room temperature at a 1:40,000-fold dilution and then the blots were washed three times with wash buffer. For chemiluminescent detection, the Supersignal West Pico Stable Peroxide solution and the Supersignal West Pico Luminol/Enhancer solution were used. Both reagents were from Pierce Chemical Company, Rockford, IL and were used as described by the manufacturer. The immunoblots were exposed to film for time frames of 5 to 30 seconds.
The *bdr* Gene Families of the Lyme Disease and Relapsing Fever Spirochetes: Potential Influence on Biology, Pathogenesis, and Evolution

David M. Roberts,* Jason A Carlyon,† Michael Theisen,‡ and Richard T. Marconi*

*Medical College of Virginia at Virginia Commonwealth University, School of Medicine, Richmond, Virginia, USA; 
†Yale School of Medicine, Yale University, New Haven, Connecticut, USA; and ‡Statens Serum Institute, Copenhagen, Denmark

Appendix

**Bacterial Cultivation, DNA Isolation, and Southern Hybridization Analyses**

Isolates belonging to the *Borrelia burgdorferi* sensu lato complex were cultivated in complete BSK-H media (Sigma) at 33°C. To cultivate the relapsing fever borreliae and other *Borrelia* species, the complete BSK-H media were supplemented with additional rabbit sera (Sigma) to a final concentration of 12% (vol/vol). Bacteria were harvested by centrifugation and washed with phosphate buffered saline (pH 7.0), and DNA was extracted (25). For Southern hybridization analyses, 5μg of DNA from each isolate was digested under standard conditions with *XbaI* and fractionated by electrophoresis in 0.8% GTG agarose gels. The DNA was transferred onto membranes for hybridization by vacuum blotting using the VacuGene system as described by the manufacturer (Pharmacia). All other Southern hybridization methods were as previously described (39).

**Immunoblot Analyses**

Bacterial cultures were grown and harvested as described above. One OD600 equivalent of cells was pelleted and resuspended in 100μl of standard SDS-sample buffer with reducing agents. The cell lysates (7μl) were fractionated by electrophoresis in 15% SDS-PAGE gels and electroblotted onto Immobilon P membranes (38). The immunoblots were blocked overnight in
blocking buffer (1X PBS, 0.2% Tween, 0.002% NaCl, and 5% nonfat dry milk) and then incubated with a 1:1,000 antisera dilutions. ImmunoPure Goat anti-mouse IgG (H+L) peroxidase conjugate served as the secondary antibody. The secondary antibody was incubated with the blots for 1 hour at room temperature at a 1:40,000-fold dilution and then the blots were washed three times with wash buffer. For chemiluminescent detection, the Supersignal West Pico Stable Peroxide solution and the Supersignal West Pico Luminol/Enhancer solution were used. Both reagents were from Pierce and were used as described by the manufacturer. The immunoblots were exposed to film for time frames of 5 to 30 seconds.