Epigenomic Landscape of Lyme Disease Spirochetes Reveals Novel Motifs

Jenny Wachter,a Craig Martens,b Kent Barbian,b Ryan O. M. Rego,c,d Patricia Rosaa

ABSTRACT Borrelia burgdorferi, the etiological agent of Lyme disease, persists in nature through an enzootic cycle consisting of a vertebrate host and an ixodes tick vector. The sequence motifs modified by two well-characterized restriction/modification (R/M) loci of B. burgdorferi type strain B31 were recently described, but the methylation profiles of other Lyme disease Borrelia bacteria have not been characterized. Here, the methylomes of B. burgdorferi type strain B31 and 7 clonal derivatives, along with B. burgdorferi N40, B. burgdorferi 297, B. burgdorferi CA-11, B. afzelii PKo, B. afzelii BO23, and B. garinii PBr, were defined through PacBio single-molecule real-time (SMRT) sequencing. This analysis revealed 9 novel sequence motifs methylated by the plasmid-encoded restriction/modification enzymes of these Borrelia strains. Furthermore, while a previous analysis of B. burgdorferi B31 revealed an epigenetic impact of methylation on the global transcriptome, the current data contradict those findings; our analyses of wild-type B. burgdorferi B31 revealed no consistent differences in gene expression among isogenic derivatives lacking one or more restriction/modification enzymes.

IMPORTANCE The principal causative agent of Lyme disease in humans in the United States is Borrelia burgdorferi, while B. burgdorferi, B. afzelii, and B. garinii, collectively members of the Borrelia burgdorferi sensu lato species complex, cause Lyme disease in Europe and Asia. Two plasmid-encoded restriction/modification systems have been shown to limit the genetic transformation of B. burgdorferi type strain B31 with foreign DNA, but little is known about the restriction/modification systems of other Lyme disease Borrelia bacteria. This paper describes the methylation motifs present on genomic DNAs of multiple B. burgdorferi, B. afzelii, and B. garinii strains. Contrary to a previous report, we did not find evidence for an epigenetic impact on gene expression by methylation. Knowledge of the motifs recognized and methylated by the restriction/modification enzymes of Lyme disease Borrelia will facilitate molecular genetic investigations of these important human pathogens. Additionally, the similar motifs methylated by orthologous restriction/modification systems of Lyme disease Borrelia bacteria and the presence of these motifs within recombinogenic loci suggest a biological role for these ubiquitous restriction/modification systems in horizontal gene transfer.

KEYWORDS Borrelia, Lyme disease, methylation, motifs, epigenetic regulation
and Asia (4, 5). The segmented genome of *B. burgdorferi sensu lato* is unstable during *in vitro* cultivation, which can lead to the loss of native linear plasmids (lp) and circular plasmids (cp). Some of the plasmids that are readily lost during culture of *B. burgdorferi* type strain B31, such as lp25 and lp28-1, are required for infectivity in a mammalian host (6–8). However, the loss of other plasmids, such as lp56, has no impact on the experimental mouse-tick infectious cycle and can aid in the genetic manipulation of the spirochete (9, 10). The enhanced transformation of spirochetes lacking lp25 and lp56 led to the discovery of the type II restriction/modification (R/M) systems encoded by the bbe02 and bbq67 loci on these plasmids, respectively (9–13). An additional R/M gene, bbh09 on lp28-3, is a homolog of bbe02 but does not appear to affect transformation efficiency (9, 13).

DNA modifications, including methylation, have diverse functions, such as impeding or aiding in the cleavage of DNA by restriction endonucleases, mismatch repair, and/or regulation of gene expression (14–16). In bacteria and archaea, methyltransferase enzymes (MTases) are responsible for methylating DNA in the form of either 6-methyladenosine (m6A), 4-methylcytosine (m4C), or 5-methylcytosine (m5C), with m6A being the most common form of methylation in prokaryotes (17, 18). Within the majority of defined bacterial methylomes, the genomic R/M motifs are almost completely methylated in the presence of the responsible R/M genes (19, 20). In fact, R/M systems are responsible for the most abundant, widely distributed DNA methylation among prokaryotes (21). Despite the high prevalence of MTases within the >200,000 bacterial and archaeal complete or draft genomes sequenced, the DNA motifs that they modify and their biological roles remain largely undefined (22–25). However, sequencing technologies capable of detecting all three major forms of bacterial DNA methylation, such as single-molecule real-time (SMRT) sequencing, have been used to generate the majority of the >2,000 mapped bacterial and archaeal methylomes (19, 23, 26–32).

Strong evidence exists for horizontal gene transfer between different strains of *B. burgdorferi sensu lato* in nature (33–38), but it is unknown if endogenous R/M genes facilitate or impede such events, particularly if heterologous R/M systems recognize and methylate different motifs. The nucleotide sequence motifs modified by the m6A activity of BBE02 and BBQ67 in *B. burgdorferi* B31 have been determined by SMRT sequencing (39), but the R/M motifs of other *B. burgdorferi sensu lato* strains remain undefined. Additionally, while strain B31 derivatives lacking bbe02 and bbq67 loci were reported to exhibit striking differences in their transcriptomic profiles relative to the wild type (wt), these comparisons were made between nonisogenic variants that differed in total plasmid contents in addition to R/M loci (39). Here, we describe the methylomes of multiple *B. burgdorferi sensu stricto*, *B. garinii*, and *B. afzelii* strains and define the novel sequence motifs modified by their respective plasmid-encoded R/M systems. Additionally, engineered deletions of R/M loci in wt B31 permitted direct comparisons of the transcriptomes of isogenic variants lacking only bbe02, only bbq67, or both loci relative to the wt clone. Contrary to a previous report of an epigenetic impact of methylation on gene expression in *B. burgdorferi* strain B31 (39), our current RNA sequencing (RNA-seq) and quantitative real-time PCR (qRT-PCR) analyses of the B31 transcriptome contradict those findings. Importantly, the identification of the motifs recognized and methylated by the R/M enzymes of Lyme disease *Borrelia* bacteria will facilitate the genetic manipulation of these important human pathogens and provide data to define a biological role for these ubiquitous R/M systems.

**RESULTS**

**Homology and REBASE searches for *Borrelia* restriction/modification genes.** Searches for BBE02/BBH09 and BBQ67 homologs were performed with whole-genome BLAST protein and nucleotide searches of *Borrelia burgdorferi sensu lato* (see Table S1 in the supplemental material [R/M homologs]) (13). These searches identified 3 R/M genes in *B. burgdorferi* N40, 3 in *B. burgdorferi* CA-11.2A, 4 in *B. burgdorferi* 297, 4 in *B. afzelii* PKo, 2 in *B. afzelii* BO23, and 2 in *B. garinii* PBr. REBASE searches were also performed to identify any additional R/M
genes; however, these generally revealed R/M enzymes that due to either size or frameshifts would not produce functional proteins (Table S1, REBASE). The phylogenetic trees were assembled from amino acid sequences identified through BLAST searches, with a greater number of BBE02/BBH09 homologs than BBQ67 homologs being identified (Fig. 1), in agreement with previous reports (13).

**Single-molecule real-time sequencing of Borrelia reveals methylation of unique nonpalindromic m6A motifs by multiple strains.** SMRT sequencing was performed to determine the motifs recognized by the R/M enzymes of 8 different *B. burgdorferi* B31 derivatives (13, 40–43), 3 different *B. burgdorferi* N40 derivatives (44), *B. burgdorferi* 297 (2), *B. burgdorferi* CA-11 (45) and clonal CA-11.2A (46), *B. garinii* PBr (47), and *B. afzelii* PKo (47) and BO23 (48), resulting in 75- to 776-fold coverage and >99.99% overall concordance with
Due to the absence of complete chromosomal sequences, contigs were utilized for chromosomal alignments of *B. burgdorferi* CA-11 and *B. afzelii* BO23. Additionally, as *B. burgdorferi* 297 does not have a sequenced chromosome, a previously sequenced and assembled chromosomal sequence of a close relative, JD1, was used for alignment. Analysis of SMRT sequencing data revealed that, for the most part, genomic DNAs (gDNAs) from different *Borrelia* strains exhibited distinct methylation motifs, although they carry homologous R/M genes (Fig. 2). The variation of motifs within the same species is consistent with the rapid evolution of the R/M system in *Borrelia*. Whether there is a significant strain-specific distribution of motifs, through analysis of the pairwise motif differences and the phylogenetic distributions, orthologous R/M enzymes methylate motifs that are more similar than those of paralogous R/M enzymes (Fig. 1 and 2). All motifs were of the m6A type, with no confindent m4C or m5C motifs being detected. Although the *Bbun40_y07* gene product of *B. burgdorferi* N40 is reported to methylate cytosine (49), cytosine methylation was not detected in either nonclonal or clonal N40 gDNA (Table 1). The number of unique methylated motifs detected for each *B. burgdorferi* B31 and N40, *B. garinii* PKo, and *B. afzelii* PBr strain was the same as the number of m6A R/M genes either known or predicted to be encoded in their genomes (Fig. 2). In contrast, *B. burgdorferi* N40, CA-11, and 297 and *B. afzelii* BO23 are predicted to harbor more R/M genes than the number of unique methylated motifs that were identified. While CA-11 harbored intact R/M genes, *Bbun40_e01* in N40 contains indels; *B. burgdorferi* 297 lacks plasmid lp28-5 carrying *Bbu297_y09* and *Bbu297_y05*, along with indels in *Bbu297_h03* and *Bbu297_j01*; and *B. afzelii* BO23 contains a frameshift in gene *BLA32_04945*. However, it is unknown if the indels in *BbuN40_e01*, *Bbu297_h03*, and

![FIG 2 Heat map depicting the methylated motifs per genome and their interpulse duration (IPD) ratios. The horizontal heat map (red) indicates the detected motifs and the pairwise motif differences (based on the Pearson correlation coefficient), with the m6A-methylated bases underlined. The vertical bar plot demonstrates the number of R/M enzymes encoded within each *Borrelia* strain and derivative, identified to the right of the heat map. Ambiguity codes for nucleotides are as follows: H is A, C, or T; K is G or T; M is A or C; N is any base; R is A or G; V is A, C, or G; and Y is C or T. Ba, *B. afzelii*; Bg, *B. garinii*; Bb, *B. burgdorferi*.

the published genome sequences (Table S2, coverage stats).
would inactivate the resulting enzymes. All other *Borrelia* strains and derivatives did not contain any indels within their respective R/M genes (Table S2, R/M sequencing). Apart from the H/WCAG motifs (the methylated adenine is underlined), clonal derivatives, nonclonal *B. burgdorferi* N40 and CA-11, and *B. afzelii* PKo had efficient methylation (>95%) of m6A motifs encoded throughout the genome, while nonclonal *B. burgdorferi* B31 M1 and 297 and *B. garinii* PBr did not methylate the majority of sites containing their respective motifs. The incomplete methylation of gDNA extracted from nonclonal populations is presumably due to the loss of plasmids encoding the R/M enzymes by a subset of the population (Table 1 and Fig. 2) (50). The two highly methylated motifs in *B. burgdorferi* B31, CGRKA modified by BBQ67 and GNAAYG modified by BB02, are consistent with those identified

| Strain or derivative | Motif | Type of methylation | No. of genomes with motif/total no. of genomes | % modified | Mean coveragef | Mean QVb |
|----------------------|-------|---------------------|--------------------------------------------|------------|----------------|----------|
| *B. burgdorferi* B31-M1 | CGRKA | m6A | 2,058/2,237 | 92.00 | 132.77 | 130.20 |
|                       | GNAAYG | m6A | 2,562/2,980 | 85.97 | 122.18 | 106.97 |
|                       | HCAAC | m6A | 10,311/16,677 | 61.83 | 148.18 | 104.20 |
| *B. burgdorferi* B31-A3 | CGRKA | m6A | 2,192/2,236 | 98.03 | 119.96 | 121.09 |
|                       | GNAAYG | m6A | 2,896/2,979 | 97.21 | 108.50 | 105.62 |
|                       | HCAAC | m6A | 7,358/16,633 | 44.23 | 145.10 | 70.37 |
| *B. burgdorferi* B31-A3-68 | GNAAYG | m6A | 2,885/2,907 | 99.24 | 119.72 | 118.81 |
|                       | HCAAC | m6A | 8,123/16,113 | 50.41 | 147.75 | 70.06 |
| *B. burgdorferi* B31-A3 Δbbo02 | CGRKA | m6A | 2,169/2,236 | 96.79 | 110.24 | 115.76 |
|                       | WCAC | m6A | 4,427/12,658 | 34.97 | 145.06 | 57.14 |
| *B. burgdorferi* B31-A3-68 Δlp56/bbb067 Δbbo02 | H/WCAAC | m6A | 5,243/16,127 | 32.51 | 162.08 | 62.15 |
| *B. burgdorferi* B31-A3 Δlp28-3/bbb09 | CGRKA | m6A | 2,195/2,198 | 99.86 | 229.90 | 282.86 |
|                       | GNAAYG | m6A | 2,940/2,944 | 99.86 | 210.64 | 225.92 |
| *B. burgdorferi* B31-A34 Tn::bbh09 | VCAAYG | m6A | 2,182/2,277 | 95.83 | 137.73 | 134.17 |
| *B. burgdorferi* N40 | VCAAYG | m6A | 2,165/2,227 | 97.22 | 234.64 | 257.67 |
| *B. burgdorferi* N40 clone 2 | VCAAYG | m6A | 2,180/2,277 | 95.74 | 218.29 | 238.61 |
| *B. burgdorferi* N40 p7 | VCAAYG | m6A | 2,499/2,535 | 98.58 | 234.85 | 306.45 |
| *B. burgdorferi* CA11 | GNAAYC | m6A | 2,504/2,535 | 98.78 | 233.08 | 303.94 |
| *B. burgdorferi* CA-11.2A clone 2 | GNAAYC | m6A | 2,222/2,581 | 86.09 | 199.87 | 101.11 |
| *B. afzelii* PKo | CTCRRA | m6A | 1,546/1,700 | 90.94 | 124.88 | 113.51 |
|                       | GGAAYC | m6A | 1,347/1,363 | 98.83 | 118.81 | 159.62 |
|                       | HMAAGG | m6A | 3,583/5,485 | 65.32 | 136.44 | 95.51 |
| *B. afzelii* BO23 | GGAAYG | m6A | 1,146/1,150 | 99.65 | 250.58 | 293.71 |
| *B. garinii* PBr | CMAAYC | m6A | 1,402/2,169 | 64.64 | 124.61 | 115.22 |
|                       | CAGC | m6A | 3,346/4,111 | 79.14 | 119.94 | 116.96 |

aThe methylated adenine is underlined.

bQV, quality value, a prediction of the error probability of a base call. Only base calls with a QV of 20 (99% accuracy) were used for analysis.

cGenomic DNA sequenced was from a nonclonal heterogeneous population.

dGenomic DNA sequenced was from a clonal population.

eB. burgdorferi* N40 clone 2 is missing lp38, and *B. burgdorferi* 297 is missing lp28-1, lp28-4, and lp28-5.

fThe mean of the total number of bases mapped to the methylated base; this includes reads from the same library molecules.

*Bbu297 j01* would inactivate the resulting enzymes. All other *Borrelia* strains and derivatives did not contain any indels within their respective R/M genes (Table S2, R/M sequencing). Apart from the H/WCAG motifs (the methylated adenine is underlined), clonal derivatives, nonclonal *B. burgdorferi* N40 and CA-11, and *B. afzelii* PKo had efficient methylation (>95%) of m6A motifs encoded throughout the genome, while nonclonal *B. burgdorferi* B31 M1 and 297 and *B. garinii* PBr did not methylate the majority of sites containing their respective motifs. The incomplete methylation of gDNA extracted from nonclonal populations is presumably due to the loss of plasmids encoding the R/M enzymes by a subset of the population (Table 1 and Fig. 2) (50). The two highly methylated motifs in *B. burgdorferi* B31, CGRKA modified by BBQ67 and GNAAYG modified by BB02, are consistent with those identified
by Casselli et al. (39). However, the H/WCAG motif of B. burgdorferi B31 has not been previously described (22, 51). The methylated motifs identified within B. burgdorferi N40, CA-11, and 297; B. garinii PBr; and B. afzelii PKo and BO23 are novel (22, 51). Further REBASE searches identified additional motifs that contain some form of the Borrelia motifs identified within this study, except for GNAAYG, GNAAYC, and HMAAG (Table S2, REBASE motifs).

The number of motifs present in Borrelia DNA is not more than would be expected by chance. To determine if individual motifs confer some type of advantage and are overrepresented in strains carrying the cognate R/M enzyme, the numbers of motifs within each genome were analyzed. This analysis assessed the numbers of motifs present in each replicon versus randomly assembled, noncoding sequences of the same length to determine if there were significantly more motifs than expected by chance in any particular replicon (Table S2, motif analysis). A false-positive “replicon” occurs if a randomly assembled sequence contains more motifs than would be expected by chance. Significance is determined by scanning primary and control sequences for each motif and computing an odds score for each position in each sequence; these scores are combined through the average odds score, and the sum test is applied to determine if the primary sequences have significantly higher scores (52). Of all the genomes analyzed, there was no evidence of selective pressure on the prevalence of R/M motifs (Table S2, motif analysis). Additionally, all genes harbored within each sequenced genome were synonymously shuffled at the third position of each codon to produce 1,000 synonymously shuffled sequences (53). Comparisons revealed that the majority of genes encoded fewer motifs than the synonymously shuffled sequences (Fig. 3).

The number of R/M motifs in the ospC gene is higher than would be expected by chance. Although there is no evidence for selective pressure on the total number of R/M motifs per replicon, we analyzed the distribution of R/M motifs in plasmid-borne genes of B. burgdorferi B31 for which there is strong evidence of horizontal gene transfer and recombination (Fig. 4 and Table S2, genes of interest) (33–38, 54–57). For most of the R/M motifs in the 7 strains analyzed, B. burgdorferi B31 ospC contained a significantly higher number than would be expected by chance. Additionally, analysis of the ospC genes of the other 6 strains revealed a significantly higher occurrence of some motifs (Table S2, ospC). In contrast, B. burgdorferi B31 ospD, dbpA, dbpB, and the erp genes revealed higher numbers of some R/M motifs than would be expected by chance, but there was also a high number (≥25%) of false-positive sequences. However, there is no appreciable difference (as determined by the log fold change [logFC]) between the numbers of motifs encoded within these genes and their synonymously shuffled counterparts (Table S2, CodonShuffle). Together, these data indicate the clustering of R/M motifs in ospC but not in other plasmid-borne genes for which there is also evidence of horizontal gene transfer, indicating an ospC-specific selective pressure.

The number of motifs and the prevalence of methylated motifs are independent of the replicon and strain, with some striking exceptions. To determine if motif prevalence varies among replicons (Fig. S2 and S3), the numbers of encoded m6A sites were compared to the length of the replicon and found to correlate (Fig. 5). Therefore, the poor correlation between the number of methylated motifs and the total number of motif sites on a replicon can be explained by either variable methylase activity or the absence of the responsible R/M locus. Analysis of methylated motifs revealed that both the CGKRA and GNAAYG motifs of B. burgdorferi B31 and the GGAYG motif of B. afzelii BO23 significantly correlate with the number of motif sites present on the replicons, regardless of the derivative being analyzed, as would be expected with methylation of ~90% of encoded motifs (Table 1 and Fig. S4A). Similarly, methylation of CTCRRA and GGAAYC of B. afzelii PKo, CA-11 and 297 significantly correlated with the number of motifs present on the replicons and methylated about 80% of their encoded motifs despite the fact that these were sequenced as nonclonal cultures (Table 1 and Fig. S4B and C). In striking contrast, however, the numbers of methylated H/WCAG motifs in B. burgdorferi B31, HMAAGG motifs in B. afzelii PKo, and CMAAYC motifs in B. garinii PBr
do not correlate with the numbers of the respective motif sites present on the replicon and were found to methylate 65% of their respective sites (Table 1 and Fig. S4). The partial methylation of gDNA by BBH09 in *B. burgdorferi* B31 likely reflects the variable activity of the enzyme in clonal B31 derivatives (A3, A3Δbbe02, A3-68, A3-68Δbbe02, and A34). However, we cannot say if the incomplete methylation of gDNA from non-clonal cultures of *B. afzelii* Pko and *B. garinii* PBr reflects the activities of the enzymes
or the loss of plasmids harboring the respective R/M loci by some spirochetes in a heterologous population (Fig. 5 and Fig. S4).

bbe02 and bbq67 transcript levels are higher in culture than in vivo. While the R/M enzymes of B. burgdorferi B31 are active in vitro, with BBE02 and BBQ67 methylating >95% of all genomic sites in clonal populations (Table 1), the transcription of these genes in vivo is unknown. Therefore, we measured bbe02 and bbq67 transcript levels in B. burgdorferi B31-A3 spirochetes grown in vitro and throughout the in vivo mouse-tick infectious cycle using qRT-PCR (Fig. 6). Transcripts for both genes were present at low levels in vitro relative to the constitutive flaB gene, and neither was detectable by qRT-PCR in total RNA isolated from infected mouse tissues. Analysis of B. burgdorferi B31-A3-infected ticks revealed low levels of bbe02 and bbq67 transcripts in larval ticks and unfed nymphs, with a transient increase in the levels of both transcripts in nymphs after feeding to repletion. Thus, bbe02 and bbq67 are expressed at very low levels throughout the mouse-tick infectious cycle, with a relative increase in transcript levels immediately following the nymphal blood meal.

Methylation by BBE02 and BBQ67 does not control the expression of genes in B. burgdorferi B31-A3. Previous analyses detailing the impact of bbq67 on shuttle vector transformation efficiency and epigenetic regulation have relied on strains lacking lp56 (9–13, 39). Therefore, deletion of the bbq67 locus was conducted to assess the effects of gene regulation on cells containing lp56. RNA-seq was performed on B. burgdorferi B31-A3 and B31-A3 derivatives containing isogenic deletions of bbe02, bbq67,
and both bbe02 and bbq67. Analysis of the RNA-seq data revealed that there was no obvious difference in gene expression regardless of the presence or absence of the R/M gene(s) when visualized with a principal-component analysis (PCA) plot to determine and visualize the maximum variation between samples (Fig. S5A). Additionally, both edgeR and DESeq2 analyses, which both normalize and analyze the data to determine differentially
expressed genes, were conducted, with similar results. A comparison of strains containing isogenic deletions of \textit{bbe02}, \textit{bbq67}, and both \textit{bbe02} and \textit{bbq67} revealed two genes that were differentially expressed. No other genes were determined to be significantly differentially expressed by edgeR and DESeq2 (Fig. 7).

As the lack of differential gene expression due to the absence of one or more \textit{R/M} genes in B31-A3 is quite different than what was reported previously by Casselli et al. (39), additional RNA-seq analyses were performed with independently prepared cDNA libraries of B31-A3, B31-A3 \( \Delta bbe02 \), B31-A3-68, B31-A3-68 \( \Delta bbe02 \), B31-A3 lp28-3\(^{-}\), B31-A34, and B31-A34 Tn\(::\)H09 (Table S1). However, similar to our previous data set, these seven strains did not reveal significant differences in gene expression that correlated with \textit{R/M} gene content. The highly passaged B31-A34 and B31-A34 Tn\(::\)H09 strains lack 9 plasmids present in the B31-A3 derivatives, causing these strains to cluster separately in the resulting PCA plot. However, the lack of \textit{bbh09} in the B31-A34 Tn\(::\)H09 strain does not result in differential gene expression compared to B31-A34, the isogenic strain from which it was derived (Fig. S5B). Additionally, edgeR and DESeq2 analyses of B31-A3 derivatives revealed that the majority of differentially expressed genes reflected the absence of lp56 in the B31-A3-68 derivatives (Fig. 8).

To confirm the RNA-seq data, qRT-PCR was performed to include differentially regulated genes with a log fold change of \(>1\). Overall, qRT-PCR could not confirm all differentially expressed genes under a less stringent log fold change of \(>1\) (Table 2, Fig. 7, and Table S2, increased and decreased expression).

In addition to the differential expression of numerous genes, Casselli et al. reported the differential expression of the alternative sigma factor \(\textit{rpoS}\) and posttranscriptional regulators of \(\textit{rpoS}\), such as \textit{bbd18}, in strains lacking \textit{bbe02} and \textit{bbq67} (39). As \textit{rpoS} is a transcription factor that regulates a large number of genes, this would have a significant

---

**FIG 6** \(\textit{bbe02}\) and \(\textit{bbq67}\) transcript levels of \textit{B. burgdorferi} B31-A3 grown in vitro or within \textit{I. scapularis}. cDNA of total RNA from \textit{B. burgdorferi} B31-A3 in vitro or within \textit{I. scapularis} at different stages underwent qPCR. \(\textit{bbe02}\) and \(\textit{bbq67}\) were expressed at lower levels within \textit{I. scapularis} than in in vitro-grown organisms, with \(\textit{bbq67}\) being present at undetectable levels in larvae and feeding nymphs. The exception is the expression of \(\textit{bbe02}\) at nymphal repletion. Significance was determined by Dunn’s multiple comparison of the Kruskal-Wallis test \((n = 3\) biological and 3 technical replicates each). ***, \(P\) value of \(<0.002\); ***, \(P\) value of \(<0.0002\).
We therefore carefully examined the expression of sigma factors and response regulators in both RNA-seq data sets and further verified the relative levels of \(rpoN\), \(rpoS\), and \(bbd18\) transcripts by qRT-PCR. Consistent with our previous analyses and different from those of Casselli et al. (39), the expression levels of sigma factors and response regulators did not differ significantly between isogenic B31 strains containing and those lacking R/M loci by either analysis (Table S2).

**DISCUSSION**

DNA modification by R/M enzymes has been recognized as having diverse impacts on microorganisms (14–16). In *B. burgdorferi* type strain B31, two plasmid-encoded type II R/M enzymes, BBE02 and BBQ67, significantly impact the stable uptake and incorporation of exogenous DNA (9–13). Analysis of 8 *B. burgdorferi* B31 derivatives...
along with *B. burgdorferi* N40, CA-11, and 297; *B. afzelii* PKo and BO23; and *B. garinii* PBr revealed 9 methylation motifs. In addition to confirming the recently described GNAAYG (modified by BBE02) and CGRKA (modified by BBQ67) motifs, an H/WCAG motif modified by BBH09 was identified within *B. burgdorferi* type strain B31 (39).

Despite the high sequence identity among BBE02 (64 to 98%) and BBQ67 (65%) orthologs analyzed in this study, the majority of methylated motifs are distinct and novel, suggesting a possible impact on genetic exchange between species and strains (Table 1 and Fig. 1 and 2). However, orthologous genes contain similar motif specificities, as observed in *B. burgdorferi* CA-11 and 297, which methylate the same motif, GNAAYC, which is similar to the *B. burgdorferi* B31 BBE02 motif (Fig. 1 and 2). The cause for motif variability, therefore, is more likely through the acquisition, loss, and pseudogenization of paralogous R/M genes than by sequence evolution among orthologs.

While the majority of genomes analyzed carry R/M genes without indels, *B. burgdorferi* 297 lacked lp28-5, which carries the R/M genes Bbu297_y09 and Bbu297_y05 and

![Volcano plots of *B. burgdorferi* B31-A3 gene expression compared to B31-A3-68 (lp56⁻) gene expression (A) and *B. burgdorferi* B31-A3 gene expression compared to B31-A3-68 Δbbe02 (Δbbe02/lp56⁻) gene expression (B). Differentially expressed genes with an FDR of <0.01 as determined by DESeq2, those with a log fold change of >2 as determined by DESeq2, and genes that have an FDR of <0.01 and a log fold change of >2 in DESeq2 are shown. Genes that have an FDR of <0.01 and a log fold change of >2 by both DESeq2 and edgeR are labeled and shown in blue.](image)

**TABLE 2** Genes with significantly decreased expression in B31-A3 Δbbe02 and B31-A3 Δbbq67 compared to B31-A3, with an FDR of <0.05 and a log₂ fold change of >1

| Gene   | Log₂ fold change | FDRᵃ      | No. of BBE02 sites | No. of BBQ67 sites | Confirmation by qPCR? |
|--------|------------------|-----------|--------------------|--------------------|-----------------------|
| bb25ᵇ  | edgeR, −3.75     | edgeR, 4.56 x 10⁻² | 0                  | 1                  | Yes, but expression does not appear to be significantly different in the absence of both bbe02 and bbq67ᵇ |
| bbb19  | edgeR, −3.15     | edgeR, 4.89 x 10⁻² | 2                  | 2                  | Did not determine as ospC expression is variable |
| bbe02  | edgeR, −2.40     | edgeR, 9.11 x 10⁻³  | 4                  | 4                  | Did not determine as bbe02 contains a deletion/insertion |
| bbh26ᵇ | edgeR, −2.21     | edgeR, 2.04 x 10⁻²  | 0                  | 0                  | Noᵇ |
| bbq67ᵇ | edgeR, −10.45    | edgeR, 1.36 x 10⁻⁵  | 2                  | 1                  | Did not determine as bbq67 is deleted |

ᵃFDR, false discovery rate, the adjusted P value that gives the proportion of false positives expected among differentially expressed genes.
ᵇSee Fig. 7.
contains several silent and missense variations within its other R/M genes, \textit{Bbu297\_h03} and \textit{Bbu297\_j01}. Additionally, all \textit{B. burgdorferi} N40 derivatives contain a single missense variation within the R/M gene \textit{Bbu40\_e01}, and BLA32_04945 in \textit{B. afzelii} BO23 contains a frameshift (see Table S2 in the supplemental material [R/M sequencing]). However, it is unknown if the identified indels would inactivate the resulting R/M enzyme. All previously and newly identified \textit{Borrelia} motifs were of the m6A type, and the number of motifs identified within each genome corresponds to the number of encoded, intact m6A R/M enzymes, except for \textit{B. burgdorferi} CA-11 (Table 1 and Fig. 2). Further work is required to assess and confirm the absence of cytosine methylation within all sequenced \textit{B. burgdorferi} N40 derivatives that carry \textit{Bbun40\_y07}, a homolog of a \textit{Haemophilus haemolyticus} cytosine methyltransferase (49).

Similar to other bacterial R/M systems, the majority of \textit{Borrelia} R/M genes analyzed in this study methylated \( \geq 95\% \) of their respective motifs (20). However, the H/WCAG motif of \textit{B. burgdorferi} B31, the GNAAYG and CGRK\_A motifs of \textit{B. burgdorferi} B31-MI, the GNAAYC motif of \textit{B. burgdorferi} 297, the HMAAAGG motif of \textit{B. afzelii} PKo, and the CMAAYC and CAGC motifs of \textit{B. garinii} PBr were not efficiently methylated. Apart from \textit{Bbu297\_h03}, \textit{Bbu297\_j01}, \textit{Bbu40\_e01}, and BLA32_04945, it is not believed that mutations have resulted in this decrease in methylation activities as all other R/M genes lack indels (Table S2, R/M sequencing). The relatively modest level of methylation (\( \sim 65 \) to 92\%) in nonclonal \textit{B. burgdorferi} B31-MI and 297, \textit{B. afzelii} PKo, and \textit{B. garinii} PBr can be explained by the absence of the plasmids carrying the R/M genes in subsets of their populations (Table 1) (41).

Successful transformation of \textit{B. burgdorferi} strain B31 derivatives largely depends on the presence or absence of R/M genes and on the transforming DNA. As previously noted, \textit{B. burgdorferi} B31 derivatives that lack one or more R/M genes, such as B31-A34 and B31-S9, are more readily transformed with shuttle vectors (9–13). Additionally, the transformation of \textit{B. burgdorferi} N40 was shown to be enhanced after the deletion of \textit{Bbun40\_e01} (65), and shuttle vector transformation has been successful in \textit{B. burgdorferi} CA-11.2A (34) and 297 (64) and \textit{B. afzelii} BO23 (66). However, the impacts of the R/M genes in \textit{B. afzelii} PKo and \textit{B. garinii} PBr on the transformation efficiency are unknown. As the relative number of \textit{B. burgdorferi} B31 CGRK\_A and GNAAYG motifs within the shuttle vectors pBSV2, pBSV2G, and pKFSS1 is consistent with previously reported transformation efficiencies (13), it is believed that the large number of \textit{B. afzelii} PKo and \textit{B. garinii} PBr motifs in these shuttle vectors would have a negative impact on the transformation efficiency.

In \textit{B. burgdorferi} B31, it has been shown that \textit{in vitro} methylation with the CpG methyltransferase M.SsI increased the rate of pBSV2 and pKFSS1 shuttle vector transformation in a BBQ67-dependent manner. Investigation of these shuttle vectors revealed that all CGRK\_A sites and \( \sim 75\% \) of GNAAYG sites contain a cytosine that would be methylated by M.SsI (12). Furthermore, previously identified sites protected from RsaI cleavage by BBQ67 methylation were found to overlap CGRK\_A methylated motifs in pBSV2, pBSV2G, and pKFSS1 (13) (Fig. S6). As the presence of R/M genes primarily affects shuttle vector transformation and has relatively little influence on the rate of allelic exchange, the construction of a shuttle vector lacking the CGRK\_A and GNAAYG motifs should allow the more efficient transformation of \textit{B. burgdorferi} B31 derivatives expressing BBQ67 and BBE02 (9–13, 67).

Our attempts to create a functional shuttle vector devoid of all CGRK\_A and GNAAYG motifs have been unsuccessful to date. The removal of these sites in the CoIE1 origin of replication or within broad-host-range plasmids such as pEP2 resulted in plasmids that were unable to replicate within \textit{Escherichia coli} (Table S2, onl). However, codon optimization and the removal of CGRK\_A and GNAAYG motifs within the \textit{flg} promoter and the anti-biotic cassettes \textit{aadA1} (GenBank accession number MW473471), \textit{accC1} (accession number MW473470), and \textit{aph} (accession number MW473472) conferred resistance to the respective antibiotics in \textit{E. coli}. Furthermore, modified \textit{flgAp-aph} was shown to confer kanamycin resistance to \textit{B. burgdorferi} B31-A3 following allelic exchange at \textit{bbq67}.

Recent work has demonstrated that a putative methyltransferase gene of \textit{Babesia bigemina}, a pathogen associated with tick vectors, is expressed only within the tick...
environment (68). Analysis of the relative levels of bbe02 and bbq67 transcripts showed significantly higher expression levels in vitro than within B. burgdorferi B31-A3-infected larvae and nymphs, except immediately following the detachment of fully engorged nymphs from naive mice (Fig. 6). However, the role of R/M gene expression within the tick, where horizontal gene transfer among Borrelia bacteria likely occurs, remains unknown (67). Analyses of B. burgdorferi B31 genes that have strong evidence for gene transfer, such as ospC (54), ospD (55), the erp genes (37, 56), and the decorin binding protein (dbp) genes (57), revealed that there is no appreciable difference in the numbers of motifs when the codons are synonymously shuffled (as determined by the log fold change [Table S2, CodonShuffle]). This may indicate that motif presence is driven by codon bias. However, a comparison of these genes with randomly assembled non-coding sequences of the same length revealed that 6 out of 15 genes and their flanking regions contained significantly more GNAAYG or CGARKA motifs than would be expected by chance (Table S2, genes of interest). ospC was the only gene analyzed to contain significantly higher numbers of both GNAAYG and CGARKA motifs than would be expected by chance. Additionally, ospC carried by the other strains in this study showed a significantly higher rate of occurrence of at least one of the newly identified motifs. Further analysis into the presence of the other known Borrelia R/M motifs in the B31 ospC gene revealed that the majority of identified motifs were also present within the central portion of the coding region. A previous investigation into recombination via gene transfer at ospC demonstrated the highest recombination breakpoints and diversity in the middle of the coding sequence (54). This leads us to speculate that the presence of Borrelia R/M motifs within ospC could facilitate the transduction of cp26 DNA or recombination via gene transfer (34, 54, 69). Therefore, while R/M motifs are typically viewed as a means to degrade foreign DNA, in the case of ospC, and perhaps other horizontally transferred DNAs, it may be beneficial by providing breakpoints for packaging DNA into phage or for recombination into the native locus. The utilization of R/M motifs for this purpose would provide a higher probability of generating novel, chimeric, immunodominant serotype-defining ospC alleles (54).

In an attempt to understand the incomplete methylation of HCAG motifs within B. burgdorferi B31, the methylation patterns of each replicon were analyzed. This revealed that the number of CGAR, GNAAYG, or HCAG motifs correlates with the length of the replicons, while the number of methylated HCAG motifs does not (Fig. 5 and Fig. S4). Neither the absence of bbe02 and bbq67 and the presence of ospC (54) can be expected by chance (Table S2, genes of interest). While this could be due to the activity of the enzyme, it could also be due to a blockage of the HCAG motif by DNA binding proteins. Of the 16,677 HCAG motifs within the B. burgdorferi B31 genome, there are 8 BosR binding sites (70–71), 2,203 EbfC binding sites (72), and 10,362 BpaB binding sites (73) within 100 bp surrounding an HCAG motif, while there are 0 and 4 BosR sites, 280 and 394 EbfC sites, and 1,506 and 1,861 BpaB sites within 100 bp of CGAR, GNAAYG motifs, respectively. While the absence of bbe02 and the presence of ospC can increase the methylation of HCAG motifs (Table S2, A3-68-LS). While this could be due to the activity of the enzyme, it could also be due to a blockage of the HCAG motif by DNA binding proteins. Of the 16,677 HCAG motifs within the B. burgdorferi B31 genome, there are 8 BosR binding sites (70–71), 2,203 EbfC binding sites (72), and 10,362 BpaB binding sites (73) within 100 bp surrounding an HCAG motif, while there are 0 and 4 BosR sites, 280 and 394 EbfC sites, and 1,506 and 1,861 BpaB sites within 100 bp of CGAR, GNAAYG motifs, respectively. Within all B31 derivatives except A34, there are 2,617 HCAG motifs that are consistently methylated and 4,682 that are never methylated. Further analysis of 100 bp around the 4,682 unmethylated motifs revealed that 1 contains the BosR binding site, 540 contain the EbfC binding site, and 2,606 contain the BpaB binding site. In comparison, there are 179 unmethylated CGARKA motifs and 418 GNAAYG motifs within B31-MI, with 0 and 1 BosR sites, 15 and 70 EbfC sites, and 103 and 221 BpaB sites within 100 bp, respectively. While this does not explain all of the unmethylated HCAG motifs, DNA binding proteins may have a negative impact on DNA methylation by BBH09.

Although the majority of methylated motifs occur fairly consistently among the replicons of each Borrelia genome analyzed, the CAAGC motif of B. garinii PBr is highly prevalent (174 sites within ~5,000 bp) and methylated at the left telomeric end of lp28-3 until it encounters the 3′ end of the bgapbr_h0006 R/M enzyme (Fig. 4 and Fig. S3). Unfortunately, as B. garinii PBr encodes two R/M enzymes, it is unknown whether this motif is methylated by BGAPBR_H0006. These motifs lie within the vlsE expression locus and vls silent cassettes of B. garinii PBr. In fact,
a closer analysis of the vls loci within *B. burgdorferi* B31, 297, and N40 and *B. afzelii* PKo revealed a high concentration of CAGC motifs within these strains and derivatives as well. As the high number of *Borrelia* R/M motifs within the *ospC* gene is presumed to aid in the horizontal transfer and antigenic variation of immunodominant OspC, the high number of CAGC motifs within the vls genes may aid in antigenic variation through gene conversion of vlsE within spirochetes.

While a previous investigation of the *B. burgdorferi* B31 methylome concluded that the BBE02 and BBQ67 R/M enzymes had a global impact on gene regulation, our current study failed to find any significant evidence for the epigenetic regulation of the *Borrelia* transcriptome by either BBE02 or BBQ67 (Table 2 and Fig. 7). Despite lowering the stringency of RNA-seq analysis to identify differentially expressed genes, these could not be consistently validated by qRT-PCR (Table S2, increased and decreased expression). Furthermore, the current study analyzed *B. burgdorferi* B31-A3, B31-A3 Δbbe02, B31-A3 Δbbq67, and B31-A3 Δbbe02 Δbbq67 that retained all native plasmids, while the previous study analyzed RNA-seq data generated from *B. burgdorferi* B31-A3 and B31-A3 Δbbe02 that lacked Ip38 and from 5A18-NP1 Δbbe02 that lacked Ip28-4 in addition to Ip56 (39). Additionally, the striking impact of R/M gene content on the expression of sigma factors and posttranscriptional regulators evidenced by Casselli et al. was not reproduced in two independently generated RNA-seq data sets during our current analysis. Variable culture conditions leading to differential sigma factor expression were most likely responsible for the putative epigenetic impact of R/M loci on gene expression in the previous study (39) and would explain this discrepancy.

**MATERIALS AND METHODS**

**Phylogenetic tree construction.** The BBE02 and BBQ67 nucleotide and amino acid sequences were used to identify similar genes within both Lyme disease and tick-borne relapsing fever spirochetes within the NCBI database as well as to search for *Borrelia* R/M genes within REBASE (see Table S1 in the supplemental material [R/M homologs and REBASE]). The identified proteins were aligned with Clustal Omega, and the percentages of identity and similarity between genes were determined with the Ident and Sim sequence analysis tool from the Sequence Manipulation Suite (74, 75). The phylogenetic tree was inferred by approximately maximum likelihood methods implemented through FastTree (76). This analysis utilized the JTT (Jones-Taylor-Thornton) model of amino acid evolution with the CAT approximation to account for the various rates of evolution across sites and estimated the local support values with the Shimodaira-Hasegawa (SH) test (77–79). The resulting tree was visualized with iTOL v 4.4.2 (80).

*8. burgdorferi* strains and growth conditions. *B. burgdorferi* strains were cultured in Barbour-Stoenner-Kelly II (BSKII) medium supplemented with 6% rabbit serum (PelFreez Biologicals, Rogers, AZ) and the appropriate antibiotics (streptomycin at 50 μg/ml and kanamycin at 200 μg/ml) at 35°C under 2.5% CO2 (81). *B. burgdorferi* B31 MI obtained from MedImmune (now AstraZeneca, Gaithersburg, MD) was subjected to PacBio sequencing as this was the gDNA used to generate the B31 genomic sequencing data for the type strain. *B. burgdorferi* strain B31-A3, an infectious clone derived from B31 MI lacking cp9, was used as the wild-type strain in this study (41). Cloning vectors were propagated using *E. coli* strain TOP10 (Invitrogen, Carlsbad, CA). All *B. burgdorferi* strains and derivatives along with plasmids utilized in this study are shown in Table S1 (bacterial strains).

**Assembly of constructs and transformation of *B. burgdorferi*.** *B. burgdorferi* was transformed by electroporation as previously described (82). Competent *B. burgdorferi* bacteria were freshly prepared from an exponential-phase culture and electroporated with 15 to 30 μg of plasmid DNA prepared from *E. coli*. Transformants were confirmed by PCR and sequencing, and the plasmid content was determined to ensure that no plasmids were lost during transformation (83). All primers used in this study are listed in Table S1 (oligonucleotides). To generate the B31-A3 Δbbq67 and B31-A3 Δbbe02 Δbbq67 strains, 500 bp upstream and downstream of bbq67 were PCR amplified and cloned flanking a kanamycin resistance cassette driven by the flg8 promoter into the pCR-Blunt II-TOPO vector (Thermo Fisher Scientific). B31-A3 and B31-A3 Δbbe02 were electroporated with the resulting construct and plated on solid medium containing either kanamycin (B31-A3 Δbbq67) or streptomycin and kanamycin (B31-A3 Δbbe02 Δbbq67) to generate the B31-A3 Δbbe02 Δbbq67 strains (Fig. 9). The antibiotic cassettes aacC1 (gentamicin 3-N-acetyltransferase), aadA1 (streptomycin 3-adenylyltransferase), and aph (aminoglycoside phosphotransferase) (kanamycin resistance) were optimized for electroporation into *B. burgdorferi* B31 through the removal of CGRKA and GNAAYG sites along with *Borrelia* codon optimization (GenScript, Piscataway, NJ). These optimized cassettes were driven by a *B. burgdorferi* R/M-minus *flg* promoter and transformed into *E. coli* TOP10 cells as part of a pUCL7 vector. These optimized antibiotic cassettes were shown to confer the appropriate antibiotic resistance (gentamicin at 5 μg/ml, spectinomycin at 100 μg/ml, and kanamycin at 50 μg/ml) in transformed *E. coli*. Additionally, *flg*-driven aph was electroporated into *B. burgdorferi* B31-A3 and conferred kanamycin resistance to transformed cells through allelic exchange at bbq67.
Genome sequencing and motif and methylation analyses. Genomic DNA (gDNA) was isolated from B. burgdorferi Mi, B31-A3, B31-A3 Δbbe02, B31-A3-68, B31-A3-68 Δbbe02, B31-A3 Δlp28-3, B31-A34, B31-A34 Trc:H90, N40, CA-11, CA-11.2a, and 297; B. garinii PBr and BO23; and B. afzelii PKo with Qiagen genomic DNA extraction according to the manufacturer’s protocol (Qiagen, Hilden, Germany). gDNA was used to generate 16 barcoded SMRTbell libraries (adapter kits 8A and 8B) and subjected to SMRT sequencing according to the manufacturer’s instructions for multiplex microbial SMRTbell libraries v2 (Pacific BioSciences, Menlo Park, CA) with the following modifications. Two pools were independently generated for sequencing on two SMRT cells on the PacBio Sequel platform (Pacific BioSciences, Menlo Park, CA). The primer and polymerase were annealed to the appropriate references using the pbsmrtpipe ds_modification_detection and sa3_ds_resequencing_fast pipelines (Pacific BioSciences). The references used for assembly were GenBank accession numbers AEO000783 to AEO000794 and AEO01575 to AEO01584 for B31; CP001651 and CP002227 to CP002242 for N40; ABJV02000001 to ABJV02000005 and CP001310 to CP001311 for PBr; CP000293 to CP0002950 for PKo; CP000163, CP000231, and CP000223 to CP0002270 for 297; ABY02000001 to ABY02000014 and CP001473 to CP001484 for CA-11; and CP018262 to CP018293 for BO23. Only base calls that had a quality value of 20 or higher were used for analysis.

To determine if the number of encoded motifs is higher than would be expected by chance, the Analysis of Motif Enrichment (AME) program of the MEME suite 5.0.5 package was utilized (52, 84). This analysis shuffles the FASTA sequences of each genome or gene to conserve 2-mer frequencies and create, at minimum, 1,000 randomized control sequences for motif comparison. AME then determines if each motif is enriched in the primary sequence using one-tailed Fisher’s exact test and performs partition maximization over all possible position-weight matrix (PWM) thresholds, with false (control sequences) and true (input sequences) positives determined by their PWM scores (52). Codons within genes were synonymously shuffled by utilizing the N3 script of CodonShuffle in which the third position of each codon was shuffled to generate 1,000 shuffled sequences for each gene (53).

The number of motifs and the number of methylated motifs per 1,000 bp were determined and mapped using the ggpubr package of the ggplot2 version 3.2.0 tool of the tidyverse package in R (86).

RNA isolation, sequencing, and quantitative real-time PCR. Total RNA was isolated from three biological replicates of B. burgdorferi strains B31-A3, B31-A3 Δbbe02, B31-A3 Δbbq67, and B31-A3 Δbbe02 Δbbq67 and from two biological replicates of B. burgdorferi B31-A3, B31-A3 Δbbe02, B31-A3-68, B31-A3-68 Δbbe02, B31-A3-68 Δlp28-3, B31-A34, and B31-A34 Trc:H90 in mid-log phase (5 × 10^5 to 9 × 10^5 cells/ml) (Table S1). Cells were collected by centrifugation and treated with RNAprotect (Qiagen). RNA was isolated using TRIzol reagent (Life Technologies, Carlsbad, CA) according to the manufacturer’s instructions and treated with 1 U of DNase I (Ambion, Foster City, CA) for 1 h at 37°C. RNA was quantified and subjected to Agilent Bioanalyzer 2200 Tape Station (Agilent, Santa Clara, CA) quality assessment. RNAs possessing RNA integrity number (RIN) values of ≥7.4 were used for downstream analysis.

cDNA libraries for RNA-seq and quantitative PCR (qPCR) were generated from these total RNAs. cDNA libraries used for RNA-seq were synthesized using Ribo-Zero and ScriptSeq complete bacterial kits with indexing primers for the synthesis of directional libraries (Illumina, San Diego, CA) according to the manufacturer’s instructions. The
quality and quantity of the resulting cDNA libraries were assessed with an Agilent DNA 1000 assay on an Agilent 2100 Bioanalyzer (Agilent) and with a Kapa library quantification kit (Kapa Biosystems, Wilmington, MA) prior to RNA-seq. RNA-seq libraries were sequenced on an Illumina MiSeq platform using v3 chemistry with 2- by 75-bp reads.

RNA-seq reads were compiled, filtered to remove any reads with PHRED scores of less than 10, and aligned to the B. burgdorferi B31 genome (GenBank accession numbers AE000783 to AE000794 and AE001575 to AE001584) using bowtie2 (87). Reads for annotated genes were determined using featureCounts (88, 89). Differential expression analysis was conducted with edgeR (90) and DESeq2 (91).

For analysis of bbd02 and bbg57 expression within the tick vector, the total RNA and resulting cDNA from B. burgdorferi B31-A3-infected ixodid ticks sampled in vitro, 1 μg of RNA was reverse transcribed using the high-capacity cDNA reverse transcriptase kit (Life Technologies) according to the manufacturer's instructions. RT reactions were performed using Q SYBR green supermix (Bio-Rad Life Sciences, Hercules, CA) with gene-specific primer sets (500 nM) (Table S1). Reactions were performed so that each experiment contained a biological and technical triplicate on a Via7 real-time PCR system (Applied Biosystems, Foster City, CA) and analyzed with PRISM software. Negative-control reactions of primers lacking a template and on RNA samples that underwent cDNA reactions in the absence of reverse transcriptase were performed with each reaction to ensure that threshold cycle (Ct) values were not obtained from primer-primer interactions or from contaminating genomic DNA. Additionally, the melt curve was analyzed for each reaction. All primers used for qPCR are listed in Table S1 (oligonucleotides).

Data availability. Sequences of standard Borrelia selectable markers that were optimized for B. burgdorferi B31 transformation by the removal of CGRKA and GNAAYG motifs are located in GenBank under accession numbers MW473470 (aacC1), MW473471 (aadA1), and MW473472 (aph). The RNA-seq data are available under GEO accession number GSE169460.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 0.2 MB.
FIG S1A to C, TIF file, 0.2 MB.
FIG S1D to G, TIF file, 0.2 MB.
FIG S3, TIF file, 1.7 MB.
FIG S4, TIF file, 0.8 MB.
FIG S5, TIF file, 0.2 MB.
FIG S6, TIF file, 0.9 MB.
TABLE S1, XLSX file, 0.1 MB.
TABLE S2, XLSX file, 0.7 MB.

ACKNOWLEDGMENTS

We thank Gang Fang, John Beaulaurier, and Pedro Oliveira for initial PacBio SMRT sequencing, analysis, and helpful discussions on Borrelia burgdorferi R/M motifs. We also acknowledge Stacy Ricklefs and Dan Bruno for their help with RNA sequencing. We are grateful to Kendal Cooper, Adam Nock, Sherwood Casjens, and Weigang Qui for reviewing the manuscript and providing helpful comments.

This research was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

REFERENCES

1. Burgdorfer W, Barbour A, Hayes S, Benach J, Grunwaldt E, Davis J. 1982. Lyme disease—a tick-borne spirochetosis? Science 216:1317–1319. https://doi.org/10.1126/science.7043737.
2. Steere AC, Grodzicki RL, Kornblatt AN, Craft JE, Barbour AG, Burgdorfer W, Schmid GP, Johnson E, Malawista SE. 1983. The spirochetological etiology of Lyme disease. N Engl J Med 308:733–740. https://doi.org/10.1056/NEJM198303313081301.
3. Lane RS, Piesman J, Burgdorfer W. 1991. Lyme borreliosis: relation of its causative agent to its vectors and hosts in North America and Europe. Annu Rev Entomol 36:587–609. https://doi.org/10.1146/annurev.en.36.010191.003103.
4. Baranton G, Assous M, Postic D. 1992. Three bacterial species associated with Lyme borreliosis. Clinical and diagnostic implications. Bull Acad Natl Sci Paris 176:1075–1095.
5. Steere AC, Strie F, Wormser GP, Hu LT, Branda JA, Hovius JWR, Li X, Mead PS. 2016. Lyme borreliosis. Nat Rev Dis Primers 2:16090. https://doi.org/10.1038/nrdp.2016.90.
6. Xu Y, Kodner C, Coleman L, Johnson RC. 1996. Correlation of plasmids with infectivity of Borrelia burgdorferi sensu stricto type strain B31. Infect Immun 64:3870–3876. https://doi.org/10.1128/IAI.64.9.3870-3876.1996.
7. Purser JE, Norris SJ. 2000. Correlation between plasmid content and infectivity in Borrelia burgdorferi. Proc Natl Acad Sci U S A 97:13865–13870. https://doi.org/10.1073/pnas.97.25.13865.
8. Labandeira-Rey M, Skare JT. 2001. Decreased infectivity in Borrelia burgdorferi strain B31 is associated with loss of linear plasmid 25 or 28-1. Infect Immun 69:446–455. https://doi.org/10.1128/IAI.69.1.446-455.2001.
9. Lawrenz MB, Kawabata H, Purser JE, Norris SJ. 2002. Decreased electropermeation efficiency in Borrelia burgdorferi containing linear plasmids Ip25 and Ip56: impact on transformation of infectious B. burgdorferi. Infect Immun 70:4796–4804. https://doi.org/10.1128/IAI.70.9.4796-4804.2002.
10. Jacobs MB, Norris SJ, Philippi-Falkenstein KM, Philipp MT. 2006. Infectivity of the highly transformable BBE02 Ip56 mutant of Borrelia burgdorferi, the Lyme disease spirochete, via ticks. Infect Immun 74:3678–3681. https://doi.org/10.1128/IAI.00043-06.
11. Kawabata H, Norris SJ, Watanabe H. 2004. BBE02 disruption mutants of Borrelia burgdorferi B31 have a highly transformable, infectious phenotype. Infect Immun 72:7147–7154. https://doi.org/10.1128/IAI.72.12.7147-7154.2004.

12. Chen Q, Fischer JR, Benoit VM, Dufour NP, Youderian P, Leong JM. 2008. In vitro Cpg methylation increases the transformation efficiency of Borrelia burgdorferi strains harboring the endogenous linear plasmid Ip36. J Bacteriol 190:7885–7891. https://doi.org/10.1128/JB.00324-08.

13. Regenstorf, Boster A, Rosa PA. 2011. Defining the plasmid-borne restriction-modification systems of the Lyme disease spirochete Borrelia burgdorferi. J Bacteriol 193:1161–1171. https://doi.org/10.1128/JB.01176-10.

14. Doerfler W. 1983. DNA methylation and gene activity. Annu Rev Biochem 52:93–124. https://doi.org/10.1146/annurev.bi.52.070183.000521.

15. Modrich P, Lahue R. 1996. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem 65:101–133. https://doi.org/10.1146/annurev.bi.65.070196.000533.

16. Casadesus J, Low D. 2006. Epigenetic gene regulation in the bacterial world. Microbiol Mol Biol Rev 70:830–856. https://doi.org/10.1128/MMBR.00016-06.

17. Korlach J, Turner SW. 2012. Going beyond five bases in DNA sequencing. Curr Opin Struct Biol 22:252–261. https://doi.org/10.1016/j.sbi.2012.04.002.

18. Beaualaurier J, Schadt EE, Fang G. 2019. Deciphering bacterial epigenomes using modern sequencing technologies. Nat Rev Genet 20:157–172. https://doi.org/10.1038/s41576-018-0081-3.

19. Murray IA, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, Packer TA, Zhai S, Dodson R, Ward MR, Ketchum KA, Dodson R, Hickey EK, Gwinn M, Dougherty B, Tomb JF, Venter JC. 1997. Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. Proc Natl Acad Sci U S A 94:13151–13156. https://doi.org/10.1073/pnas.94.24.13151.

20. Oliveira PH, Ribas JW, Garrett EM, Trzilova D, Kim A, Sekulovic O, Mead EA, Luong K, Morgan RD, Boitano M, Anton BP, Murray IA. 2011. De novo discovery of DNA methylation during single-molecule real-time sequencing. Nucleic Acids Res 39:13077–13087. https://doi.org/10.1093/nar/gkt1201.

21. Furuta Y, Namba-Fukuyo H, Shibata TF, Nishigou S, Suzuki Y, Sugano S, Hasebe M, Kobayashi I. 2014. Methylation diversification through changes in DNA methylation transferase sequence specificity. PLoS Genet 10:e1004272. https://doi.org/10.1371/journal.pgen.1004272.

22. Luch-Senar M, Luong K, Lorén-Rico V, Delgado J, Fang G, Spittle K, Turner SW, Korlach J, Serrano L. 2013. Comprehensive methylation characterization of Mycoplasma genitalium and Mycoplasma pneumoniae at single-base resolution. PLoS Genet 9:e1003191. https://doi.org/10.1371/journal.pgen.1003191.

23. Qiu W-G, Schutzer SE, Bruno JF, Attie O, Xu Y, Dunn JJ, Fraser CM, Cañas J, Sr. Luft BJ. 2004. Genomic exchange and plasmid transfers in Borrelia burgdorferi sensu stricto revealed by three-way genome comparisons and multilocus sequence typing. Proc Natl Acad Sci U S A 101:14150–14155. https://doi.org/10.1073/pnas.0402745101.

24. Engbers CH, Kimmel BJ, Bones JL, Elas AF, Rosa P, Samuels DS. 2001. Transduction by φBB-1, a bacteriophage of Borrelia burgdorferi. J Bacteriol 183:4771–4778. https://doi.org/10.1128/JB.183.11.4771-4778.2001.

25. Resnick J, Vargas LC, Morganod EF, Xue V, Hernandez Y, Pagan P, Fraser LM, Liddig C, Schutzer SE, Luft BJ, Cañas J, Sr. Qiu W-G. 2011. Pervasive recombination and synaptic genome diversification driven by frequency-dependent selection in Borrelia burgdorferi, the Lyme disease bacterium. Genetics 189:951–966. https://doi.org/10.1373/genetics.111.130773.

26. Rosa PA, Schwab T, Hogan D. 1992. Recombination between genes encoding major outer surface proteins A and B of Borrelia burgdorferi. Mol Microbiol 6:3031–3040. https://doi.org/10.1111/j.1365-2958.1992.tb01761.x.

27. Stevenson B, Cañas J, Rosa P. 1998. Evidence of past recombination events among the genes encoding the Esp antigens of Borrelia burgdorferi. Microbiology (Reading) 144:1869–1879. https://doi.org/10.1042/MB980121.

28. Lively I, Gibb's CP, Schuster R, Dorner F. 1995. Evidence for lateral transfer and recombination in OsC variation in Lyme disease bacteria. Mol Biol Evol 12:768–776. https://doi.org/10.1093/mbie/12.4.768.

29. Cassell T, Touraud Y, Scheidegger A, Arnold WK, Proux A, Stevenson B, Briisbe CA. 2018. DNA methylation by restriction modification systems affects the global transcriptome profile in Borrelia burgdorferi. J Bacteriol 200:e00395-18. https://doi.org/10.1128/JB.00395-18.

30. Fraser CM, Cañas J, Schwab T, Hogan D. 2002. Clonal polymorphism of Borrelia burgdorferi strain B31 MI: implications for mutagenesis in an infectious strain background. Infect Immun 70:2139–2150. https://doi.org/10.1128/IAI.70.4.2139-2150.2002.

31. Blomberg SA, Webster DR, Lee JE, Travers KJ, Olofse E, Clark TA, Korlach J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods 7:461–465. https://doi.org/10.1038/nmeth.1459.
45. Schwan TG, Schumpp ME, Kastens RH, Clover JR, Wong J, Daugherty M, Strickler M, Rosa PA. 1993. Distribution and molecular analysis of Lyme disease spirochetes, Borrelia burgdorferi, isolated from ticks throughout California. J Clin Microbiol 31:3096–3108. https://doi.org/10.1128/JCM.31.12.3096-3108.1993.

46. Margolis N, Rosa PA. 1993. Regulation of expression of major outer surface proteins in Borrelia burgdorferi. Infect Immun 61:2207–2210. https://doi.org/10.1128/IAI.61.5.2207-2210.1993.

47. Caimano MJ, Iyer R, Eggers CH, Gonzalez C, Morton EA, Gilbert MA, Bezier J, Bartlett F, Mann J, Schendel J, Olson K, Feil EJ, Tarr CD, Friedrich C, Kado CIJ. 2013. Molecular analysis of the ospC and dpa genes regulated by mammalian host-specific signals. Infect Immun 71:3371–3383. https://doi.org/10.1128/IAI.71.6.3371-3383.2003.

48. Hübner A, Yang X, Nolen DM, Popova TG, Cabello FC, Norgard MV. 2001. Expression of Borrelia burgdorferi OspC and DpbA is controlled by a RpoN-RpoS regulatory pathway. Proc Natl Acad Sci U S A 98:12724–12729. https://doi.org/10.1073/pnas.231442998.

49. Chan K, Alter L, Barthold SW, Parveen N. 2015. Disruption of bbe02 by insertion of a luciferase gene increases transformation efficiency of Borrelia burgdorferi and allows live imaging in Lyme disease susceptible C3H mice. PLoS One 10:e0129532. https://doi.org/10.1371/journal.pone.0129532.

50. Grimm D, Elias AF, Tilly K, Rosa PA. 2003. Plasmid stability during in vitro growth of Borrelia burgdorferi BB256. Plasmid 50:64-71. https://doi.org/10.1016/S0148-2184(03)00087-2.

51. Roberts RJ, Vincze T, Posfai J, Macelis D. 2010. REBASE: a uniﬁed framework and an evaluation on ChIP data. BMC Bioinformatics 11:165. https://doi.org/10.1186/1471-2105-11-165.

52. McLeay RC, Bailey TL. 2010. Motif enrichment analysis: a uniﬁed framework and an evaluation on ChIP data. BMC Bioinformatics 11:165. https://doi.org/10.1186/1471-2105-11-165.

53. Schulte-Spechtel U, Fingerle V, Goettner G, Rogge S, Wilske B. 2006. Function during the enzootic cycle. Mol Microbiol 65:1193–1205. https://doi.org/10.1111/j.1365-2958.2006.05860.x.

54. Barbour AG, Travinsky B. 2010. Evolution and distribution of the ospC gene of Borrelia burgdorferi sensu lato grande. J Mol Evol 70:329–342. https://doi.org/10.1007/s00239-009-2482-x.

55. Marconi RT, Samuels DS, Landry RK, Garon CF. 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. https://doi.org/10.1128/jb.176.15.4572-4582.1994.

56. Stevenson B, Miller JC. 2003. Intra- and interbacterial genetic exchange of Borrelia spirochete efB genes. J Mol Evol 57:309–324. https://doi.org/10.1007/s00239-003-4282-x.

57. Schulte-Spechtel U, Fingerle V, Goettner G, Rogge S, Wilske B. 2006. Molecular analysis of decorin-binding protein A (DpaB) reveals five major groups among European Borrelia burgdorferi sensu lato strains with impact for the development of serological assays and indicates lateral gene transfer of the dpaB gene. Int J Med Microbiol 290(Suppl 40):250–266. https://doi.org/10.1016/j.ijmm.2006.01.006.

58. Yang X, Alani SM, Norgard MV. 2003. The response regulator RpII is essential for the expression of major membrane lipoproteins in Borrelia burgdorferi. Proc Natl Acad Sci U S A 100:11001–11006. https://doi.org/10.1073/pnas.1834315100.

59. Caimano MJ, Wyer R, Eggers CH, Gonzalez C, Morton EA, Gilbert MA, Schwartz I, Radolf JD. 2007. Analysis of the RpoS region in Borrelia burgdorferi in response to mammalian host signals provides insight into RpoS function during the enzootic cycle. Mol Microbiol 65:1193–1217. https://doi.org/10.1111/j.1365-2958.2007.05860.x.

60. Caimano MJ, Eggers CH, Hazelzet KRO, Radolf JD. 2004. RpII is not central to the general stress response in Borrelia burgdorferi but does control expression of one or more essential virulence determinants. Infect Immun 72:6433–6445. https://doi.org/10.1128/IAI.72.11.6433-6445.2004.

61. Caimano MJ, Eggers CH, Gonzalez CA, Radolf JD. 2005. Alternate sigma factor RpoS is required for the in vivo-specific repression of Borrelia burgdorferi plasmid Ip54-borne ospA and Ip66 genes. J Bacteriol 187:7845–7852. https://doi.org/10.1128/JB.187.15.7845-7852.2005.

62. Fisher MA, Grimm D, Henion AK, Elias AF, Stewart PE, Rosa PA, Gherardini FC. 2005. Borrelia burgdorferi ori(S) is required for mammalian infection and vector transmission but not for tick colonization. Proc Natl Acad Sci U S A 102:5162–5167. https://doi.org/10.1073/pnas.0408536102.

63. Brooks CS, Hefry PS, Jolliff SE, Akins DR. 2003. Global analysis of Borrelia burgdorferi genes regulated by mammalian host-specific signals. Infect Immun 71:3371–3383. https://doi.org/10.1128/IAI.71.6.3371-3383.2003.

64. Geissler K, Carkhuff MA, Grimm D, Henion AK, Elias AF, Stewart PE, Rosa PA, Gherardini FC. 2005. Borrelia burgdorferi ori(S) is required for mammalian infection and vector transmission but not for tick colonization. Proc Natl Acad Sci U S A 102:5162–5167. https://doi.org/10.1073/pnas.0408536102.

65. Barbour AG, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16:1114–1116. https://doi.org/10.1093/oxfordjournals.molbev.a026201.

66. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259. https://doi.org/10.1093/nar/gkyz239.

67. Barbour AG. 1984. Isolation and cultivation of Lyme disease spirochetes. Yale J Biol Med 57:521–525.
82. Samuels DS, Mach KE, Garon CF. 1994. Genetic transformation of the Lyme disease agent Borrelia burgdorferi with coumarin-resistant gyrB. J Bacteriol 176:6045–6049. https://doi.org/10.1128/jb.176.19.6045-6049.1994.

83. Norris SJ, Howell JK, Odeh EA, Lin T, Gao L, Edmondson DG. 2011. High-throughput plasmid content analysis of Borrelia burgdorferi B31 by using Luminex multiplex technology. Appl Environ Microbiol 77:1483–1492. https://doi.org/10.1128/AEM.01877-10.

84. Bailey TL, Bodén M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208. https://doi.org/10.1093/nar/gkp335.

85. Krzywinski M, Schein J, Briel I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645. https://doi.org/10.1101/gr.092759.109.

86. Wickham H. 2016. ggplot2: elegant graphics for data analysis, 1st ed, vol 8. Springer-Verlag, New York, NY.

87. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

88. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930. https://doi.org/10.1093/bioinformatics/btt656.

89. Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res 47:e47. https://doi.org/10.1093/nar/gkz114.

90. McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 40:4288–4297. https://doi.org/10.1093/nar/gks042.

91. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8.

92. Savage CR, Jutras BL, Bestor A, Tilly K, Rosa P, Tourand Y, Stewart PE, Brissette CA, Stevenson B. 2018. Borrelia burgdorferi SpoVG DNA- and RNA-binding protein modulates the physiology of the Lyme disease spirochete. J Bacteriol 200:e00033-18. https://doi.org/10.1128/JB.00033-18.

93. Jutras BL, Savage CR, Arnold WK, Lethbridge KG, Carroll DW, Tilly K, Bestor A, Zhu H, Seshu J, Zückert WR, Stewart PE, Rosa PA, Brissette CA, Stevenson B. 2019. The Lyme disease spirochete's BpuR DNA/RNA-binding protein is differentially expressed during the mammal-tick infectious cycle, which affects translation of the SodA superoxide dismutase. Mol Microbiol 112:973–991. https://doi.org/10.1111/mmi.14336.