DNA Methylation by Restriction Modification Systems Affects the Global Transcriptome Profile in *Borrelia burgdorferi*

Timothey Casselli
University of North Dakota

Yvonne Tourand
University of North Dakota

Adam Scheidegger
University of North Dakota

William K. Arnold
Addgene

Anna Proulx
University of North Dakota

DNA Methylation by Restriction Modification Systems Affects the Global Transcriptome Profile in *Borrelia burgdorferi*

Authors

Timothey Casselli, Yvonne Tourand, Adam Scheidegger, William K. Arnold, Anna Proulx, Brian Stevenson, and Catherine A. Brissette

DNA Methylation by Restriction Modification Systems Affects the Global Transcriptome Profile in *Borrelia burgdorferi*

Timothy Casselli,[a] Yvonne Tourand,[a] Adam Scheidegger,[a] William K. Arnold,[b]* Anna Proulx,[a] Brian Stevenson,[b] Catherine A. Brissette[a]

[a]Department of Biomedical Sciences, University of North Dakota, School of Medicine and Health Sciences, Grand Forks, North Dakota, USA

[b]Department of Microbiology, Immunology and Molecular Genetics, University of Kentucky, School of Medicine, Lexington, Kentucky, USA

ABSTRACT Prokaryote restriction modification (RM) systems serve to protect bacteria from potentially detrimental foreign DNA. Recent evidence suggests that DNA methylation by the methyltransferase (MTase) components of RM systems can also have effects on transcriptome profiles. The type strain of the causative agent of Lyme disease, *Borrelia burgdorferi* B31, possesses two RM systems with *N*6-methyladenosine (m6A) MTase activity, which are encoded by the *bbe02* gene located on linear plasmid lp25 and *bbq67* on lp56. The specific recognition and/or methylation sequences had not been identified for either of these *B. burgdorferi* MTases, and it was not previously known whether these RM systems influence transcript levels. In the current study, single-molecule real-time sequencing was utilized to map genome-wide m6A sites and to identify consensus modified motifs in wild-type *B. burgdorferi* as well as MTase mutants lacking either the *bbe02* gene alone or both *bbe02* and *bbq67* genes. Four novel conserved m6A motifs were identified and were fully attributable to the presence of specific MTases. Whole-genome transcriptome changes were observed in conjunction with the loss of MTase enzymes, indicating that DNA methylation by the RM systems has effects on gene expression. Genes with altered transcription in MTase mutants include those involved in vertebrate host colonization (e.g., *rpoS* regulon) and acquisition by/transmission from the tick vector (e.g., *rrp1* and *pdeB*). The results of this study provide a comprehensive view of the DNA methylation pattern in *B. burgdorferi*, and the accompanying gene expression profiles add to the emerging body of research on RM systems and gene regulation in bacteria.

IMPORTANCE Lyme disease is the most prevalent vector-borne disease in North America and is classified by the Centers for Disease Control and Prevention (CDC) as an emerging infectious disease with an expanding geographical area of occurrence. Previous studies have shown that the causative bacterium, *Borrelia burgdorferi*, methylates its genome using restriction modification systems that enable the distinction from foreign DNA. Although much research has focused on the regulation of gene expression in *B. burgdorferi*, the effect of DNA methylation on gene regulation has not been evaluated. The current study characterizes the patterns of DNA methylation by restriction modification systems in *B. burgdorferi* and evaluates the resulting effects on gene regulation in this important pathogen.

KEYWORDS *Borrelia burgdorferi*, DNA methylation, Lyme disease, epigenetics, methyltransferase, regulation of gene expression, restriction modification, single-molecule real-time sequencing

Address correspondence to Timothy Casselli, timothy.casselli@med.und.edu.

* Present address: William K. Arnold, Addgene, Cambridge, Massachusetts, USA.

Lyme disease is the most prevalent vector-borne disease in North America (1, 2). Clinical manifestations can include arthritis, carditis, and neurological complications such as meningitis, cranial and peripheral neuritis, facial nerve palsy, and cognitive decline (3). This multisystem disease is caused by infection with a genetically heterogeneous group of spirochetes, including *Borrelia burgdorferi* and related species (4). Both inter- and intraspecies phenotypic variation has been reported within this genospecies complex, and attempts have been made to correlate this diversity to genotypic differences (5–15). The contribution of alternative drivers of phenotypic diversity in *B. burgdorferi*, such as the epigenetic regulation of gene expression, has not been evaluated.

DNA methylation is the product of methyltransferase enzymes (MTases) and has been described in both prokaryotic and eukaryotic organisms (16, 17). In prokaryotes, MTases are best understood as a component of restriction modification (RM) systems (18). RM systems are nearly ubiquitous in prokaryotes (19) and play a pivotal role in the defense against foreign DNA. Specifically, type II RM systems harbor both an MTase enzyme that recognizes and methylates a specific DNA sequence and a restriction endonuclease (RE) enzyme that recognizes and cleaves at or near the unmethylated form of the recognition sequence (20). MTase and RE functions can be carried out by two separate proteins or can exist as two distinct functional domains of a single protein (i.e., subtype IIC). As DNA methylation protects from cleavage by the RE enzyme, the specific recognition motif is modified at ~100% of the sites in the genome. In contrast to RM systems, orphan MTases that lack cognate RE enzymes are less common in prokaryotic genomes and have roles in cell cycle regulation, phase variation, and the regulation of gene expression (21, 22).

The predominant form of DNA methylation in prokaryotes is *N6*-methyladenosine (m6A) (23), and two distinct m6A MTases have been described in *B. burgdorferi* (24, 25). Clones of the type strain B31 lacking plasmid lp25 (lp25$^-$) or lp56 (lp56$^-$) are much more competent when transformed with exogenous DNA but not with *B. burgdorferi*-derived DNA, suggesting that these two genetic elements harbor RM systems (24, 25). The predicted amino acid sequence similarity implicated two genes encoding putative bifunctional MTase-RE proteins belonging to the PD-(D/E)XK superfamily of type IIC restriction enzymes: *bbe02* on lp25 and *bbq67* on lp56 (24). Inactivation of *bbe02* replicates the transformation phenotype of lp25$^-$ strains, confirming this gene as the lp25-resident restriction endonuclease (26). Additionally, Southwestern blot analysis of *B. burgdorferi* genomic DNA shows decreased levels of m6A in strains lacking either *bbe02* or lp56, supporting the bifunctional nature of these proteins (25).

Although the effects of *bbe02* and *bbq67* on transformation efficiency have been demonstrated, the specific recognition and/or methylation sequence motifs have not been identified for either of the *B. burgdorferi* MTases. It is not currently known whether the MTases associated with these RM systems can also function as gene expression regulators. The complements of identified MTases differ between commonly studied *B. burgdorferi* isolates (25), and mutant strains lacking endogenous MTases are often used as surrogates for "wild-type" strains in laboratory studies because of their increased transformation efficiency and ease of genetic manipulation (27–33). As such, a more comprehensive characterization of the DNA methylation systems in *B. burgdorferi* and their effects on gene regulation will aid in the interpretation of these studies and could prove crucial for a more thorough understanding of *B. burgdorferi* biology.

In the present study, single-molecule real-time (SMRT) sequencing (34–36) was utilized to map the genome-wide DNA methylation pattern and to identify consensus modified motifs in wild-type *B. burgdorferi* as well as MTase mutants lacking either *bbe02* alone or both *bbe02* and the lp56 plasmid that harbors the *bbq67* gene. Four novel conserved m6A modification motifs were identified and were fully attributable to the presence of either *bbe02* or lp56 (*bbq67*). Next, the transcriptome profiles were compared between strains to determine the effects of altered global DNA methylation patterns on gene expression. Whole-genome transcriptome changes were observed in conjunction with the loss of MTase enzymes, including changes in genes involved in

the colonization of vertebrate hosts and acquisition/transmission by the tick vector, indicating that DNA methylation by RM systems has effects on gene expression in *B. burgdorferi*. The results of this study provide a comprehensive view of the DNA methylation profile in *B. burgdorferi*, and the integrated gene expression profiles add to the relatively new body of research on gene expression consequences resulting from differential genome methylation by RM systems in bacteria.

RESULTS AND DISCUSSION

Single-molecule real-time sequencing reveals conserved nonpalindromic m6A motifs. DNA methylation was assayed in genomic DNA isolated from *Borrelia burgdorferi* B31 (see Materials and Methods for strain information) using SMRT sequencing (Pacific Biosciences). An initial analysis using the resequencing pipeline of SMRT analysis tools revealed variable sequence coverage from 151-fold to 428-fold among plasmids, with 969-fold coverage of the chromosome (mean, 690-fold across the entire genome) and >99.98% overall concordance with the published genome, with no significant insertions/deletions (see Table S1 in the supplemental material). Due to the high level of coverage and concordance, the published genome was used for all further analysis to enable easier comparisons with previously reported data sets. In addition to base sequence information, the presence of DNA modifications, including methylation, is measurable using SMRT sequencing. The kinetic characteristics of DNA polymerization, such as the duration between successive base incorporations (termed interpulse duration [IPD]), are altered by the presence of a modified base in the DNA template. DNA methylation is characterized by a detectable increase in the time lapse of base incorporation compared with the expected duration (IPD ratio) (34–36). The phred-like modification quality value (QV) is assigned to each base using information on coverage and consistency of IPD ratios and indicates the confidence in methylation calls from the assay. SMRT sequencing detected predominantly modified adenine bases in *B. burgdorferi*, as shown in a plot of modification QV against sequencing coverage (Fig. 1A, red dots), matching the modification profile for m6A methylation. This finding is consistent with the predicted m6A MTase function of the RM systems in *B. burgdorferi* as well as with a previous report demonstrating the presence of m6A via a Southwestern blot (25). An analysis of the sequences surrounding m6A sites revealed four conserved m6A modification motifs with a total of 5,606 m6A sites in the genome (Fig. 1B). All four motifs are novel, in that they have not been described for any previously characterized MTase (37). Only two of the motifs (CGRKA and GNAAYG) were modified at ~100% of available sites, consistent with RM protection from cleavage by a cognate RE enzyme. The remaining two motifs (DGDAAGG and DGGCATG) were modified at 40% and 66% of available sites, respectively; a hallmark of orphan m6A modification not associated with restriction protection. This inefficient modification was reflected by a reduction in mean IPD ratios compared to those of fully methylated motifs (Fig. 1D). An alignment of position weight matrices (Fig. 1C) revealed conserved guanine residues at positions −3 and +2 surrounding the methylated adenines between the partially methylated DGDAAGG and DGGCATG motifs and the fully methylated GNAAYG motif, suggesting these motifs may represent promiscuous methylation by a single MTase. The mean coverage depth for the four described motifs ranged from 352-fold (GNAAYG) to 430-fold (DGGCATG), well above the recommended 25× to 100× coverage for base modification detection. To ensure that the motifs were not simply technical artifacts due to oversequencing, the modification motifs were reanalyzed on a subset (20%) of the zero-mode waveguides (ZMWs) from the SMRT cell. All four motifs were identified at between 75× to 93× coverage in the low-coverage reanalysis, supporting their methylation in *B. burgdorferi* (see Fig. S1 for comparison of low-coverage and high-coverage position weight matrices and coverage statistics).

All m6A motifs identified in *B. burgdorferi* were found to be nonpalindromic (Fig. 1), indicative of the recognition sequences of type IIS restriction endonucleases (e.g., FokI) (20). Type IIS REs cleave DNA at a specific location downstream of short nonpalindromic recognition sequences. Additionally, all motifs identified in this study were methylated
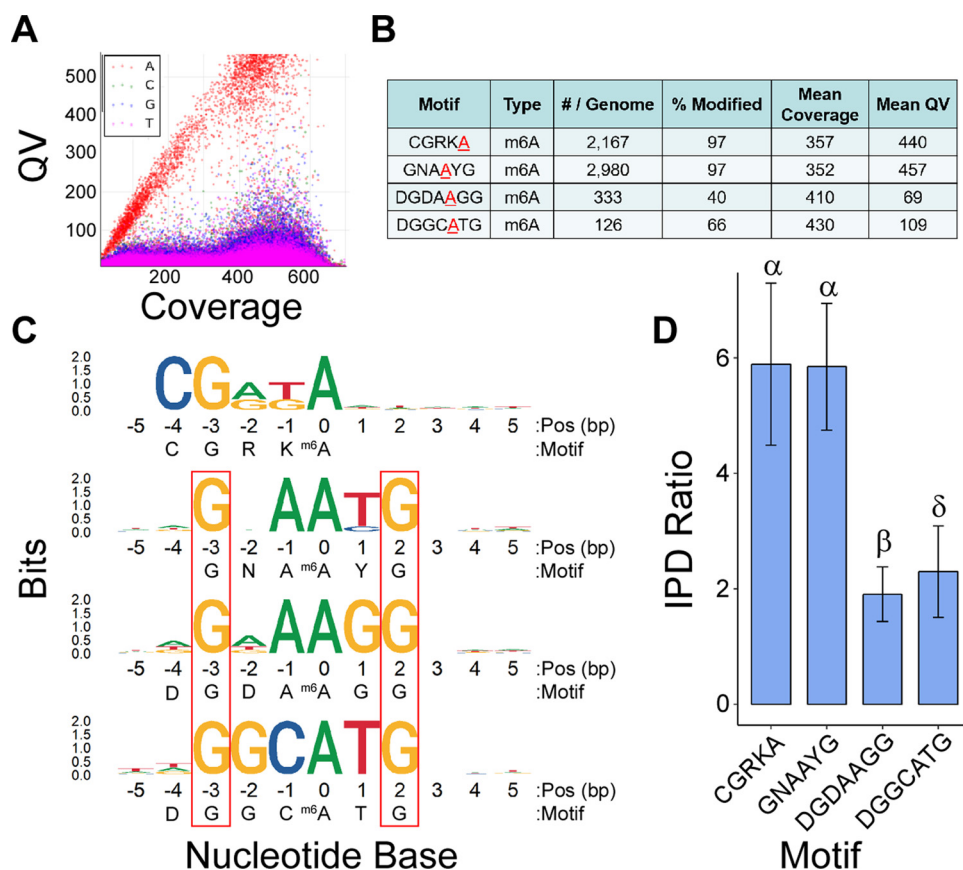
FIG 1 DNA modification in *B. burgdorferi* is predominantly m6A and is located within conserved motifs. (A) Dot plot demonstrating per-base modification scores as a measure of sequencing coverage. Clustering of red dots shows adenine as the predominant modified base detected by SMRT sequencing in *B. burgdorferi*. (B) Detected conserved m6A motifs along with the relative abundances and percentages modified for each motif. Average sequencing coverage and modification scores are also shown. (C) Sequence logo position weight matrices showing identified conserved modification motifs, with the modified adenine aligned at base position 0. Frequency of base occurrence at a given position is denoted by the height of each letter. Red boxes indicate conserved sites for three of the motifs, suggesting potential promiscuous modification by a single enzyme. (D) Mean IPD ratios ± standard deviations (SDs) for all modified bases in each motif. Greek letters indicate significantly different groups as determined by one-way analysis of variance (ANOVA) followed by a Holm-Sidak analysis of multiple comparisons ($P < 0.001$). Decreased mean IPD ratios for DGDAAG and DGGCATG reflect incomplete modification at these sites.

on a single strand within the motif. While the majority of RM systems methylate both strands of the DNA molecule, a novel class of RM systems, termed MmeI family RM enzymes (type IIG), was recently found to possess type IIS restriction specificity, in which the MTase function modifies a single strand within the recognition sequence (38, 39). This single-stranded methylation by type IIG enzymes is sufficient for restriction protection, likely due to the requirement for multiple unmethylated motifs on a single molecule for efficient cleavage (38).

Methylated adenines are not evenly distributed throughout the *B. burgdorferi* genome. To evaluate the distribution of m6A sites throughout the genome, the numbers of modification motifs were determined per 1,000-bp region. As shown in Fig. 2A and B, there was variability in the frequencies of m6A sites both between and within genomic elements. The chromosome had an increased median number of m6As per 1,000 bp compared to that of the plasmids, with the exception of cp26 and lp28-2, and the median numbers per region also differed between different plasmids as determined by nonoverlapping 95% confidence intervals. Each genomic element typically contained at least one outlier region of a higher or lower frequency of modifications. Interestingly, the center of the chromosome contained a cluster of increased m6A outliers (Fig. 2B) that mapped closely to the reversal in GC skew (data not shown),
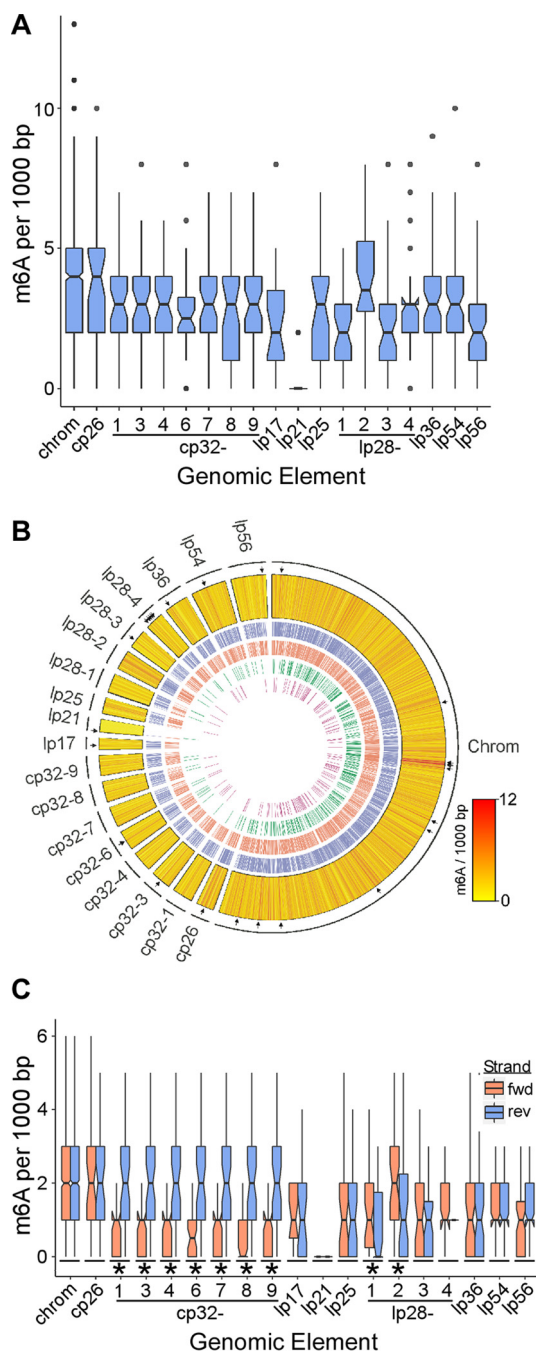
FIG 2 m6A distribution is not uniform across the genome. (A) Distributions of total m6A sites per 1,000 bp for each genomic element. Notches represent 95% confidence intervals (CI) of the medians as determined by medians $\pm$ 1.58 $\times$ IQR/($n^{1/2}$). Nonoverlapping CIs demonstrate variation in methylation frequency between elements. Dots represent outlier regions within each element. (B) Circos plot showing heatmap of modifications per 1,000 bp (yellow→red; outer track). The scale bar indicates the number of m6As per region. Black tick marks outside the heat map show locations of outlier regions from panel A. Inner tracks show individual motif locations (GNAAYG, blue; CGRKA, red; DGDAAGG, green, DGGCATG, purple). (C) Per-strand distribution of m6A sites per 1,000 bp for each genomic element. Notches represent 95% CIs of the medians. Asterisks below boxes represent genomic elements with nonoverlapping 95% CIs between forward and reverse strands. Note that outliers are not displayed.

suggesting a possible role in DNA replication (40); however, this pattern was not consistent for any of the plasmids presumed to have similar replication mechanisms. Strikingly, the plasmid lp21 contained just four conserved m6A motifs (Fig. 2A), all within the first 3,074 bp of the "left" end (Fig. 2B). This plasmid contains a 63-bp tandem

TABLE 1 Number of *B. burgdorferi* genes containing m6A motifs

Location	No. modified/total no. (%)[a]	Median no. of m6As/gene (95% CI)	Maximum no. of m6As/gene
Gene bodies	1,349/1,784 (76)	3 (2.83–3.17)	33
Upstream region[b]	758/1,784 (42)	1 (0.94–1.06)	7

[a]Number of regions containing at least one m6A motif.
[b]Modifications found in 200-bp region located 5′ of the start site.

repeat element that spans ~11 kb (41). As the repeated element does not contain any conserved sites for m6A modification, lp21 contains very few modified bases overall.

Methylation of all four identified m6A motifs was found to be localized to a single strand within the motif. The plotting of the m6A motifs per 1,000-bp fragment by strand (forward versus reverse) revealed a strong reverse-strand bias for all of the cp32 plasmids as well as less striking forward-strand biases for lp28-1 and lp28-2. The cp32 plasmids are stable prophages (42–44) that contain an average strand bias for harbored genes of ~15:1 that is not seen in the rest of the genome (e.g., chromosome coding strand bias of 1:1), suggesting a possible coding strand bias for m6A motifs. Separate analyses of the distributions of forward/reverse strand m6A sites did not reveal any obvious patterns with implications for DNA replication (i.e., leading/lagging strand bias) (data not shown).

Of the 1,784 genes and noncoding RNAs (ncRNAs) annotated in or reported for *B. burgdorferi* (45), 76% contained at least one m6A motif within the gene body (Table 1; see also Table S2 for complete gene body m6A counts). The median number of m6A sites per gene was three, with a maximum of 33. An examination of 200-bp regions upstream of these genes that may contain promoter and/or 5′-untranscribed-region (UTR) elements revealed the 46% possessed at least one modification motif, with a median of 1 and a maximum of 7 (Table 1; see also Table S3 for complete upstream region m6A counts).

All four conserved motifs showed a significant bias toward the sense coding strand of genes when the total number of gene-body m6A sites was considered ($P < 0.002$) (Fig. 3A), whereas the GNA<u>A</u>YG motif appeared to contribute most to this bias on a per-gene basis (Fig. 3B). This bias toward the sense strand may be a mechanism to maximize genomic DNA methylation events while minimizing the effects on gene expression. Modified bases on the template DNA strand (i.e., the antisense strand) may slow the rate of transcription elongation, much like the premise for SMRT sequencing described for DNA polymerase. This has been described for various DNA modifications, including m6A in eukaryotic systems (46, 47); however, the effects of m6A modification on prokaryote RNA polymerase kinetics are not known. Nonetheless, it is possible that the overrepresentation of m6A modifications on the sense strand of genes minimizes any potential effects on RNA polymerase kinetics in *B. burgdorferi*.

Hot spots of DNA modifications surrounding and within gene bodies can provide clues as to potential functional consequences. A metagene analysis of m6A distributions within genes did not reveal any obvious bias of m6A sites toward the 5′ or 3′ ends of the genes or clustering at putative promoter elements; however, there were decreases in the m6A frequency at both the gene start and end sites (Fig. 3C). Unsupervised clustering with different parameters did not reveal any reliably reproducible subgroups with patterns of m6A enrichment across gene bodies or at the gene start/end sites (data not shown).

m6A motifs detected by SMRT sequencing are attributable to *bbe02* and *bbq67*. The genome sequence of *B. burgdorferi* encodes three intact putative bifunctional MTase-RE proteins belonging to the PD-(D/E)K superfamily of type IIC restriction enzymes: the products of *bbe02* on lp25, *bbq67* on lp56, and *bbh09* on lp28-3 (24, 48). Of these, only the products of *bbe02* and *bbq67* have been demonstrated to contribute to restriction protection against exogenous DNA and to the presence of genomic m6A (24, 25). To assign recognition sequence specificity to individual MTases, the methyl-
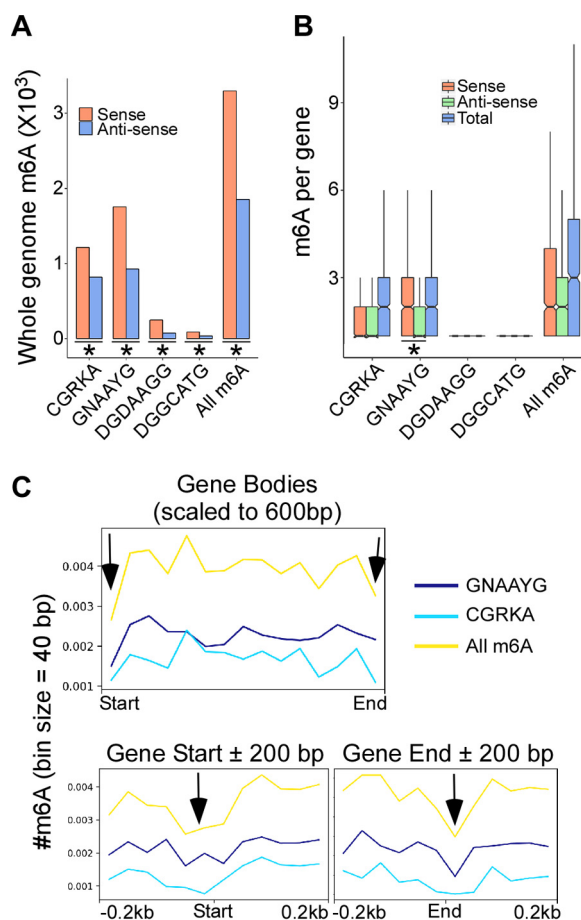
FIG 3 Gene body m6A shows a coding strand bias. (A) Total gene body m6A counts by coding strand (sense versus antisense). All motifs showed a significant sense strand bias as determined by Fisher's exact test ($P < 0.002$), as denoted by asterisks below bars. (B) Distribution of m6A counts per gene by coding strand. The asterisk indicates a sense strand bias for the GNAAYG motif as determined by nonoverlapping CIs between sense and antisense distributions. Outliers are not displayed. (C) Metagene analysis showing average m6A per 40-bp bin size across all gene bodies scaled to 600 bp (top panel), as well as at gene start and end sites ± 200 bp (bottom panels). Arrows show an apparent reduction in m6A frequency near the start and end coordinates.

omes of mutant strains lacking *bbe02* alone or both *bbe02* and the linear plasmid lp56, which harbors the *bbq67* gene, were assayed using SMRT sequencing (see Materials and Methods for strain information).

B. burgdorferi B31 clone 5A18 was previously characterized to have lost lp56 (49). An isogenic mutant of clone 5A18 (5A18-NP1, referred to as strain BbΔe02/q67⁻) was also previously generated to contain an insertional inactivation in the *bbe02* gene using a kanamycin resistance cassette (*kan*) (26). To generate a mutant strain lacking *bbe02* alone, the NP1 mutation was introduced into wild-type *B. burgdorferi* as described in Materials and Methods (A3-NP1, further referred to as strain BbΔe02). Wild-type and mutant strains were screened by PCR for the presence of MTase genes (Fig. 4A). As expected, the template DNA from the BbΔe02/q67⁻ mutant did not produce a PCR product with primers specific for lp56. Unexpectedly, primers specific for the deleted region of *bbe02* yielded amplicons from all strains. As *bbh09* and *bbe02* are homologues with 89% nucleotide identity (48), this may have been the result of nonspecific amplification of *bbh09*. To resolve this issue, NheI-digested plasmid DNA was analyzed by Southern blotting and probed for the deleted region of *bbe02* as well as the *kan* insertion marker (Fig. 4B). As expected, only DNA from the wild-type strain was positive for *bbe02* and negative for the *kan* marker, whereas both BbΔe02 and BbΔe02/q67⁻ strains were negative for *bbe02* and positive for *kan*. Additionally, a positive band was
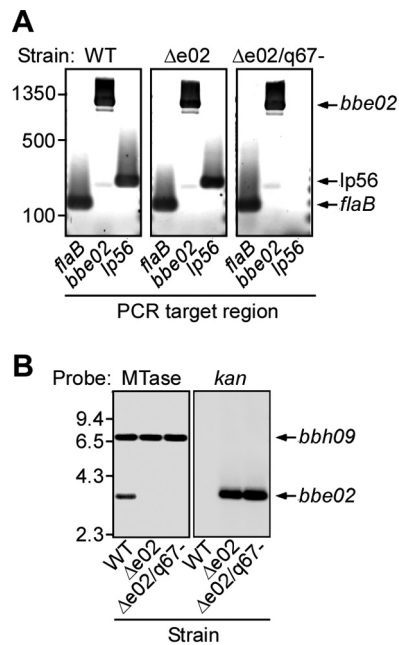
FIG 4 Confirmation of MTase gene disruption mutants. (A) PCR analysis of total DNA from the wild type (WT) and the deletion mutants as shown. Primer pairs for the chromosomal *flaB* gene were used as a positive control. All strains gave positive signals for the *bbe02* gene, likely due to nonspecific amplification of the homologous *bbh09* gene. As expected, no signal was detected for plasmid lp56 in the double MTase mutant. (B) Southern blot analysis of NheI-digested plasmid DNA from wild-type and deletion mutant strains probed for *bbe02* (MTase) and *kan* as shown. Mutant strains were positive for the *kan* gene and demonstrated the expected loss of signal for the *bbe02* gene, while the homologous *bbh09* gene was not affected. Locations of size standards are shown to the left of the blots (bp for PCR, kbp for Southern blot), while expected sizes of the target fragments are shown on the right.

seen for all three strains at the expected size for the fragment containing the *bbh09* gene when probed for *bbe02*. Collectively, these data confirm the specific deletion of *bbe02* in both mutants, without a disruption of the homologous *bbh09* locus.

SMRT sequencing of genomic DNA from BbΔe02 and BbΔe02/q67⁻ strains revealed unambiguous changes in the profiles of conserved m6A motifs for each strain (Table 2). The BbΔe02 mutant retained m6A modification at 95% of CGRKA motifs. However, none of the other motifs found in wild-type *B. burgdorferi* were detected in this strain, identifying GNAAYG as the primary methylation target of BBE02, with DGDAAGG and DGGCATG as promiscuous yet inefficient noncanonical target sites of this enzyme. No conserved m6A motifs were identified in the BbΔe02/q67⁻ mutant, implicating CGRKA as the target site for methylation by BBQ67.

No evidence was found of conserved m6A motifs in the absence of *bbe02* and *bbq67* in the given *in vitro* system, despite the presence of the intact *bbh09* gene and other loci with putative MTase motifs, as predicted by ReBase (37), as well as with evidence for an orphan MTase in relapsing fever *Borrelia* (50). A potential explanation for this phenomenon is that BBH09 is relevant exclusively during the enzootic cycle. It has been

TABLE 2 Conserved m6A motifs in MTase mutants

Motif[a]	Type	No./genome	% modified		
			WT	Δe02	Δe02/q67⁻
CGRKA	m6A	2,980	97	95	NA[b]
GNAAYG	m6A	2,167	97	NA	NA
DGDAAGG	m6A	333	40	NA	NA
DGGCATG	m6A	126	66	NA	NA

[a]Methylated adenines are underlined.
[b]NA, not detected by SMRT analysis.

noted that although the predicted amino acid sequences of BBE02 and BBH09 share 92% similarity, the isoelectric points of these homologs are quite different (6.99 versus 8.07, respectively), which may affect their activities (24). It is therefore feasible that standard *in vitro* culture conditions for *B. burgdorferi* favor the enzymatic activity of BBE02 but not BBH09. Likewise, sequence differences between BBH09 and BBE02 may reflect differences in substrate bases. To maximize the accuracy of modification scores in the current study, the observed IPD ratios were compared to the expected signatures of only m6A and *N*4-methylcytosine (m4C); thus, other types of modification may not be detected (36, 51). Finally, *bbh09* may not encode a functional MTase and may be somehow required as an accessory molecule for BBE02 or BBQ67 functionality. RM enzymes typically function as dimers or tetramers and often associate as heteromers (52). The requirement for accessory proteins for BBE02 and BBQ67 functionality is unlikely, however, given that clones lacking lp28-3 and other plasmids are not enriched during mutagenesis of *B. burgdorferi* as is often seen for lp25 and lp56 harboring the *bbe02* and *bbq67* genes (24, 25) (unpublished observations).

Altogether, the data presented here on methylation motif specificity and previous data on restriction protection demonstrate that *bbe02* and *bbq67* both encode bifunctional RM enzymes with structures similar to those of type IIC endonucleases and with MTase sequence specificities similar to those of type IIG enzymes (38). This family typically recognizes 6- to 7-nucleotide-long contiguous sequence motifs and modifies only a single strand for restriction protection. They possess both endonuclease and DNA methyltransferase activities in the same polypeptide and require *S*-adenosylmethionine (AdoMet) for endonuclease activity. Additionally, the restriction function of type IIG enzymes cleave ~20 nucleotides (nt) downstream of the recognition sequence and require interaction between two molecules bound at specific recognition sites to achieve cutting. Although the BBE02- and BBQ67-specific motifs were not typically found in close proximity, it remains to be determined whether more distant DNA-protein complexes associate to catalyze activity.

Deletion of RM systems results in global changes in gene expression in *B. burgdorferi*. Although the primary function of RM systems in prokaryotes is thought to be for restriction protection from foreign DNA, recent studies have demonstrated global effects on gene expression by DNA methylation from RM systems (53, 54). Additionally, orphan MTase enzymes with phase variability have been shown to affect the global transcriptome (55). To examine this possibility in *B. burgdorferi*, transcriptome profiles were generated by deep sequencing of rRNA-depleted total RNA isolated from the wild type and the MTase mutants. In all, 1,317 annotated genes and ncRNAs (45) were detected. Using a false discovery rate (FDR) of ≤0.01, 417 genes were differentially expressed in the BbΔe02 strain compared to that in the wild type, whereas 564 genes were differentially expressed in the BbΔe02/q67⁻ strain (see Tables S4 and S5 for complete expression profiles). Figure 5A and B shows the number, direction, and fold change of genes differentially expressed in either the BbΔe02 or BbΔe02/q67⁻ strain, as well as the considerable overlap (305 genes) of those differentially expressed in both MTase mutants. Clusters of Orthologous Groups (COG) functional categories were all similarly represented in the differentially expressed genes lists, with the majority (58%) of differentially expressed genes unclassified (data not shown). These data suggest that the DNA methylation patterns by both BBE02 and BBQ67 have distinct and widespread effects on gene expression in *B. burgdorferi*.

The enzootic cycle of *B. burgdorferi* requires the transmission of the bacterium between animal hosts (predominantly small mammals) via ticks from the genus *Ixodes* (56). This dual-host lifestyle consisting of tick acquisition/colonization, transmission/acute mammalian infection, and persistent mammalian infection requires distinct gene expression patterns between these vastly different environments. Thus, much attention has been given to elucidating the regulators of adaptive gene expression profiles (57, 58). Figure 5C shows the direction and magnitude of expression changes for the selected genes involved in regulating gene expression in each MTase mutant. The
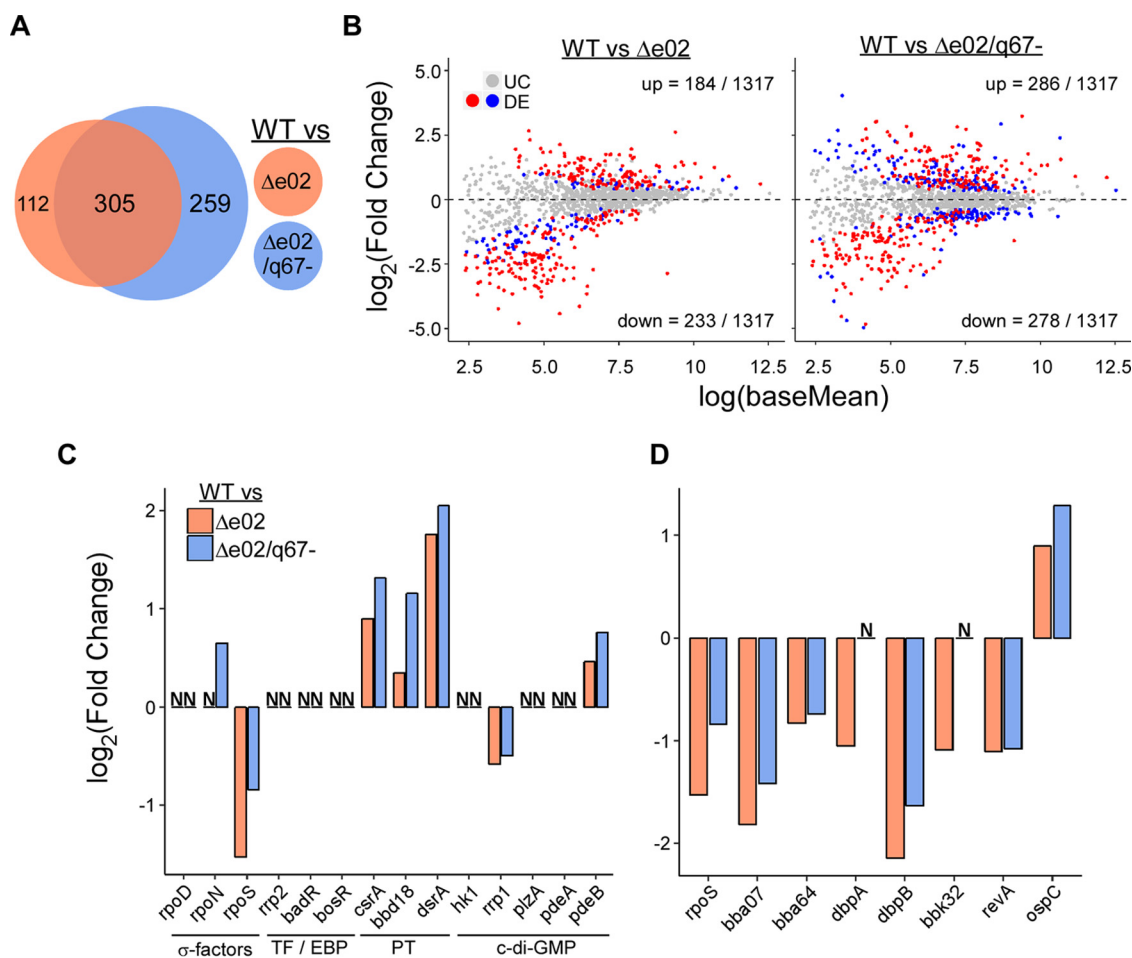
FIG 5 Deletion of MTases results in genome-wide changes in gene expression. (A) Venn diagram showing the number of unique and shared differentially expressed genes (FDR ≤ 0.01) in MTase mutants compared to that in the wild type (WT) as determined by RNA-seq. (B) Fold change versus expression strength for all detectable genes. Red dots represent differentially expressed (DE) genes common to both MTase mutants, whereas blue dots represent genes differentially expressed in either the BbΔe02 or BbΔe02/q67⁻ strain alone. Gray dots represent genes not significantly different between WT and MTase mutants (UC). Numbers of significantly upregulated (up) and downregulated (down) genes are shown as proportions of all detectable genes. (C) Differential expression of selected previously identified gene regulators. The alternative sigma factor *rpoS* and posttranscriptional regulators of *rpoS* were most affected. N, no significant differential expression; TF, transcription factor; EBP, enhancer-like binding protein; PT, posttranscriptional regulator; c-di-GMP, cyclic di-GMP second messenger synthesis and/or effector pathway. (D) Differential expression of *rpoS* and selected *rpoS*-dependent genes. All genes followed the expected pattern of differential expression based on *rpoS* expression with the exception of *ospC*, suggesting dysregulation by some additional regulatory network for that gene.

housekeeping RNA polymerase sigma factor *rpoD* was unaffected in either of the mutants. In contrast, the sigma factor *rpoS*, which is required for colonizing the mammalian host, was downregulated in both the BbΔe02 and BbΔe02/q67⁻ strains. This was coupled with a concomitant increase in the expression of the posttranscriptional *rpoS* repressors *csrA* and *bbd18* and the small regulatory RNA *dsrA* (29, 59, 60). Likewise, several RpoS-dependent and other effector genes involved in host colonization showed a concurrent downregulation, including *bba07*, *bba64*, *dbpA*, *dbpB*, *bbk32*, and *revA* (61) (Fig. 5D). Interestingly, the "canonical" RpoS-dependent gene *ospC* was upregulated in both MTase mutants. The *ospC* gene contains a unique operator region located immediately upstream of its RpoS-dependent promoter that is hypothesized to allow input from additional gene regulatory networks in order to independently repress *ospC* expression while continuing to express the RpoS regulon during persistent infection (62, 63). Thus, it is possible that any decrease in *ospC* expression in the MTase mutants due to the downregulation of *rpoS* is offset by an even larger derepression or activation by dysregulation of some yet-unidentified regulatory network. Notably, the

magnitude of *rpoS* downregulation was smaller in the BbΔe02/q67⁻ strain than in the BbΔe02 strain. This pattern was consistent across all *rpoS*-dependent genes, including *ospC* (Fig. 5D).

In addition to dysregulation of the *rpoS* regulon and genes involved in mammalian host colonization, both the BbΔe02 and BbΔe02/q67⁻ mutants showed significant decreases in the expression of the *rrp1* gene, which encodes a diguanylate cyclase response regulator (64), coupled with increases in the expression of the phospho-diesterase-encoding *pdeB* (65). Together, these data suggest the potential for decreased availability of cyclic di-GMP within MTase mutants. Cyclic di-GMP has been demonstrated to be important during the acquisition and transmission of *B. burgdorferi* by ticks and has been shown to have interplay with the *rpoS* regulon (66, 67). Collectively, these results demonstrate that gene regulatory networks relevant to survival during the enzootic cycle are influenced by endogenous RM systems in *B. burgdorferi*.

To examine the association between methylation events and disrupted gene expression profiles, the distributions of m6A motifs were compared between "unchanged" and "differentially expressed" genes for each mutant strain. No methylation profile was overrepresented in the list of differentially expressed genes compared to the proportion of all genes (Fig. 6A). Additionally, no association was observed when unchanged and differentially expressed genes were compared with respect to the presence/absence of m6A sites located immediately upstream of the genes (Fig. 6C) or with respect to the cumulative m6A distributions across/surrounding gene bodies (Fig. 6D) or transcription start sites (68) using metagene analysis (data not shown). Curiously, the median number of *bbe02* modification motifs was higher in those gene bodies where expression levels were unaffected than in differentially expressed genes in both mutants (Fig. 6B). It is likely that only a subset of m6A modifications have meaningful implications for gene expression changes, and those changes can be amplified through the affected gene regulatory networks. Therefore, a global analysis of m6A distributions between unchanged and differentially expressed genes reveals little insight into the specific m6A changes responsible for altering the transcriptome. Similarly, Fang et al. previously reported that the expression of more than one-third of the genes in *Escherichia coli* strain C227-11 were significantly altered when the EcoGIII RM system was deleted, although there was not a compelling correlation detected between m6A modification events and differentially expressed genes (53).

One caveat of the differential gene expression data reported here is the potential for confounding effects from genetic differences other than those in MTase genes. Although the BbΔe02 strain is an isogenic deletion mutant of the wild-type strain, the BbΔe02/q67⁻ mutant lacks the entire lp56 plasmid harboring the *bbq67* gene and has some other plasmid differences compared to the other two strains. Although it cannot be ruled out that some of the gene expression differences observed in the BbΔe02/q67⁻ strain are due to differences in its plasmid profile, it is likely that the vast majority of differentially expressed genes are a result of the lack of methylated DNA motifs for three reasons: (i) the differences in gene content in the BbΔe02/q67⁻ strain do not include any predicted gene expression regulators, (ii) the large numbers of affected genes in the isogenic BbΔe02 mutant demonstrate the global effects of differential genome methylation on gene expression, and (iii) despite differences in clonal origin between the two mutant strains, there is considerable overlap in the gene expression differences compared to the wild-type strain. Although this study clearly demonstrates the general phenomenon of altered gene expression resulting from differentially methylated genomes in *B. burgdorferi*, specific individual differences need to be verified experimentally in future investigations.

Despite the global changes in gene expression observed in the MTase mutants when cultivated *in vitro*, it should be noted that *B. burgdorferi* strains lacking either *bbe02* alone or *bbe02* and *bbq67* are fully infectious in a laboratory murine model of infection as determined by the 50% infective dose (ID$_{50}$), joint swelling, and histopathology (26). These strains are also capable of completing the tick-mouse enzootic cycle
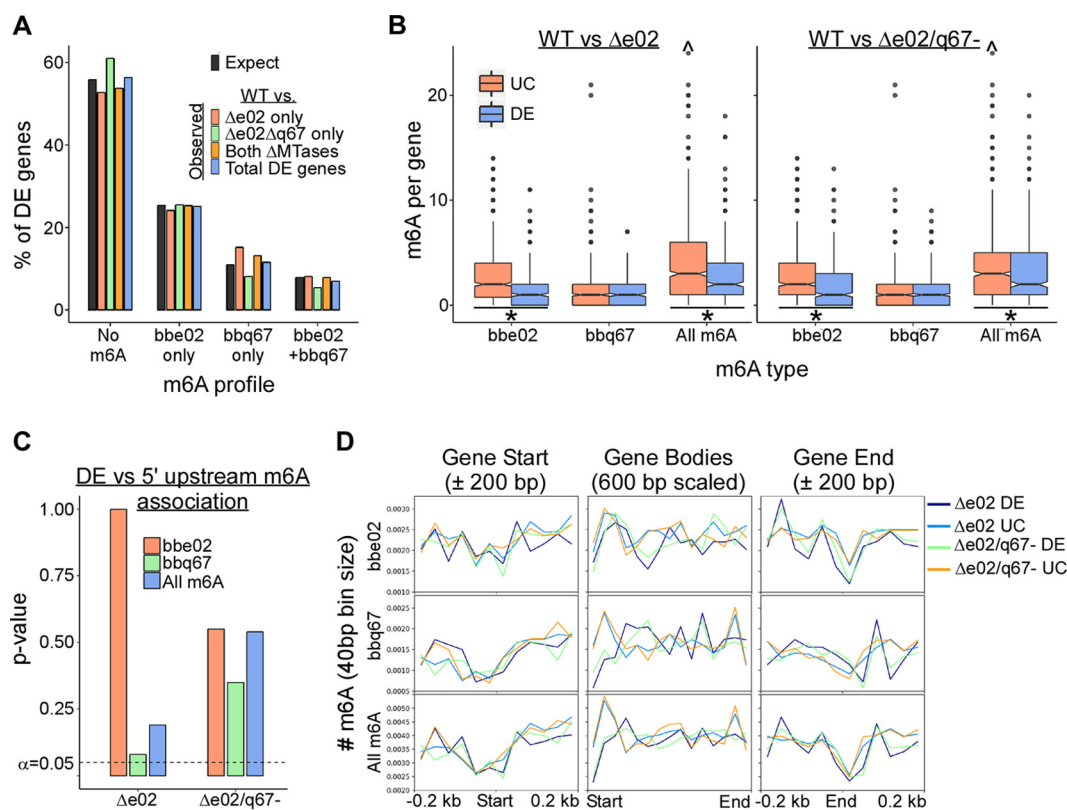
FIG 6 Correlation between m6A profile and differential expression. (A) Proportion of genes differentially expressed (DE) in MTase mutants by methylation status. Black bars show the expected percentages based on the proportion of all gene bodies containing either no m6A motifs (no m6A), at least one modification site for *bbe02* alone (*bbe02* only) or *bbq67* alone (*bbq67* only), or at least one motif for both MTases (*bbe02* + *bbq67*). Colored bars show the observed percentages of differentially expressed genes with the corresponding m6A profiles as indicated in the legend. No methylation profile was overrepresented in the population of DE genes for either MTase mutant. (B) Number of m6A sites per gene body by differential expression status for each MTase mutant. The number of *bbe02* motifs per gene was lower in differentially expressed (DE) genes compared to that for unchanged (UC) genes in both MTase mutants as determined by nonoverlapping 95% CIs and denoted by asterisks below the boxes. ^, outlier found beyond the *y* axis maximum. (C) Association between the presence of at least one m6A modification located within 200 bp upstream of a gene start site and differential expression status. Fisher's exact test was performed on 2-by-2 contingency tables comparing gene counts for methylation status (+ or −) for each MTase with differential expression status (DE or UC). Bar plot shows *P* values from these tests for each m6A profile grouped by mutant strain. No association was below the α value of 0.05 threshold for significance, indicating a lack of association between upstream m6A and differential expression. (D) Metagene analysis by differential expression status showing average m6A per 40-bp bin size across all gene bodies scaled to 600 bp (middle panel) and at gene start and end sites ± 200 bp (left and right panels, respectively). Colored lines show comparisons of differentially expressed (DE) and unchanged (UC) genes in each mutant strain. All four groups showed similar profiles with no apparent change in the overall magnitudes and no significant shift from the gene ends.

under laboratory conditions, though lower pathogen burdens were reported in ticks for a strain lacking lp56 containing the *bbq67* gene (69). As the changes observed here for many genes, including *rpoS*, were relatively modest compared to the magnitude of induction in response to host-specific signals (70), it is possible that host signals are able to overcome gene dysregulation by producing levels sufficient to establish infection in a laboratory setting. Nonetheless, the loss of DNA methylation may have more subtle implications for gene regulation patterns and pathogen fitness during the enzootic cycle than those detected in previous studies.

Conclusions. The data presented here characterize for the first time the global m6A methylome during *in vitro* cultivation of *B. burgdorferi* and assign specific methyltransferases responsible for all detectable conserved modification motifs. The MTases BBE02 and BBQ67 both methylate their primary target sequences at ~100% efficiency, consistent with restriction protection. Additionally, BBE02 exhibits inefficient, promiscuous, noncanonical activity at two additional sequence motifs that are not likely to provide restriction protection due to the much lower efficiency of methylation. It is not

known whether these additional motifs carry biological significance or if they are simply coincidental off-target methylation events. Nonetheless, as a consequence of the introduction of >5,000 methyl groups to the genome, methylation by both MTases has widespread effects on gene expression.

It has been well documented that temporal differences in DNA methylation by orphan MTases with differential/phase variable expression can act as a means of adaptive gene expression with phenotypic effects, including altered virulence (71, 72). Conversely, intact RM systems that serve to protect from restriction cleavage are not typically thought to demonstrate such temporal differences in DNA methylation efficiencies, as this would leave the genome susceptible to degradation by the bacterium's own restriction enzymes. As such, the effects of intact RM systems on global gene expression profiles have largely been ignored. Nonetheless, gene regulatory networks evolve in the context of a bacterium's endogenous RM systems, and changes in DNA methylation patterns through alterations in the complement/specificity of these RM systems may likewise be expected to have effects on transcription. Recently, another type IIG MTase in *Campylobacter jejuni* was found to affect the expression of more than 200 genes between 1.5- and 5-fold, similar to the effect sizes seen for both *bbe02* and *bbq67* (54). These and the findings reported here provide evidence that DNA methylation by intact type IIG RM systems have modest yet widespread effects on gene regulatory networks, albeit with smaller effect sizes than those reported for orphan MTases such as dam in *Salmonella* (55).

Homologues of the *bbe02* gene have been found in all identified *B. burgdorferi* isolates to date (25). In contrast, the presence of *bbq67* homologues and other predicted MTases varies between isolates. Rego et al. hypothesized that these differences may act to drive strain heterogeneity in *B. burgdorferi* and suggest a mechanism primarily related to restriction protection from horizontal gene transfer (25). The data presented here support alterations in gene expression profiles as another potential mechanism driving strain heterogeneity through RM diversification. Given the observed effects of altered DNA methylation on genes involved in host colonization and vector acquisition/transmission, changes in the complement of RM systems could have strain-specific effects on host range or virulence.

The data presented in this study have implications for laboratory models of Lyme disease. The MTase genes *bbe02* and *bbq67* were first identified due to the high transformability of strains lacking these loci (24). As a consequence, many laboratories utilize MTase-deficient strains, including the BbΔe02/q67⁻ strain used in this study (5A18-NP1), as surrogates for wild-type *B. burgdorferi* because of their ease of genetic manipulation (27–33). The findings reported here suggest that these strains may not be appropriate model organisms for precise studies of gene regulatory networks in *B. burgdorferi*, particularly when interpreting conclusions for strains with differential MTase enzyme profiles. There is a clear need for improvements to these model systems, including *B. burgdorferi*-specific cloning vectors devoid of recognition sites for cleavage by BBE02 or BBQ67 or mutant strains lacking the restriction endonuclease function of these RM enzymes while retaining MTase functions. Such tools would serve to overcome the low transformation efficiency of wild-type *B. burgdorferi* while retaining the native methylome and gene expression profiles.

Overall, this study provides the first evidence of the functional consequences of RM systems in *B. burgdorferi* beyond restriction protection. Further studies involving strains with different profiles of RM systems, as well as biochemical characterizations of RM enzymes, will further our understanding of the biology of this important group of human pathogens.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The *B. burgdorferi* strains A3 (wild type) and 5A18-NP1 (BbΔe02/q67⁻) were kindly provided by Patti Rosa and Steven Norris, respectively. Both are clones of the sequenced B31 isolate (48), and their respective plasmid profiles have been previously described (26, 73). A confirmatory analysis of plasmid content revealed that our stock of A3 was lacking plasmid lp38 (not shown). A3 was used to generate the mutant A3-NP1 (BbΔe02) strain as described. All *B. burgdorferi*

clones were grown at 35°C under 5% CO_2 in modified Barbour-Stoenner-Kelly medium (BSK) supplemented with 6% rabbit serum (74). Mutant strains were grown with kanamycin (200 μg/ml). Cell densities and growth phases were monitored by visualization under dark-field microscopy and by counting using a Petroff-Hausser counting chamber.

Generation of the A3-NP1 mutant strain. To introduce the NP1 mutation into a wild-type A3 background, plasmid DNA was isolated from *B. burgdorferi* 5A18-NP1 with a Plasmid Midi kit (Qiagen, Valencia, CA) and used as the PCR template. A region encompassing the deleted portion of the *bbe02* gene containing an integrated kanamycin resistance cassette was amplified using primers P16 and P17 (26) (see Table S6 in the supplemental material) and subsequently cloned into the vector pJET1.2 (Thermo Fisher Scientific, Waltham, MA). The resulting deletion plasmid was transformed into an *E. coli* intermediate strain for maintenance and propagation, and the insert was confirmed by restriction digest and DNA sequencing analysis.

A3 electrocompetent cells were cultivated and prepared as previously described (75). The cells were transformed with 50 μg of the purified deletion construct described above, after which, transformations were recovered for 24 h in BSK followed by plating by limiting dilution with kanamycin selection to isolate clonal transformants. The deletion mutants were initially identified by PCR screening for the kanamycin resistance gene using primers P8 and P9 (76) (Table S6). Isolated plasmid DNA from kanamycin-positive clones was screened by PCR to determine endogenous plasmid content using primers specific for regions unique to each plasmid, as described previously (77). DNA from one representative clone containing all parental endogenous plasmids and the inserted kanamycin gene was selected for further analysis by Southern blotting as described below.

Southern blot hybridization. Plasmid DNA isolated from wild-type, BbΔe02, and BbΔe02/q67⁻ strains was digested with NheI (New England BioLabs, Ipswich, MA) and cleaned with a PCR cleanup kit (Qiagen). Digested DNA was separated by 1% agarose gel electrophoresis followed by bidirectional transfer to two Amersham Hybond-N+ membranes (GE Healthcare, Piscataway, NJ) for Southern blot analysis. The probes for *kan* and *bbe02* were generated by PCR from a purified deletion construct plasmid and A3 plasmid DNA, respectively, using primers P8 and P9 (*kan*) or P143 and P144 (*bbe02*) (Table S6) with the DIG Probe synthesis kit (Roche, Indianapolis, IN), according to the manufacturer's instructions. Bands were detected using anti-DIG Fab fragment conjugated to alkaline phosphatase and visualized using the chemiluminescent substrate CDP-Star (Roche).

Genome sequencing and methylation analysis. DNA was isolated from a pool of three independent late-log-phase cultures (5 ml per culture) for each *B. burgdorferi* strain by standard phenol-chloroform extraction, concentrated by precipitation with isopropanol, and cleaned with Agencourt AMPure XP beads (Beckman Coulter). Purified DNA was outsourced to the deep-sequencing core at the University of Massachusetts Medical School for library preparation and sequencing. Libraries were generated using a 20-kb shear protocol, loaded onto one SMRT cell per sample, and sequenced on a PacBio RSII instrument using P6-C4 sequencing chemistry and a 360-min data collection protocol.

Sequencing coverage/concordance was evaluated by aligning subreads to the published *B. burgdorferi* B31 genome (RefSeq numbers AE000783 to AE000794 and AE001575 to AE001584) using the Resequencing analysis tool from the SMRT analysis package 2.3.0 (http://www.pacb.com/devnet/). The detection of modified bases and clustering of methylated sites to identify methylation-associated motifs were performed with the RS_Modification_and_Motif_analysis tool. IPD ratios (observed/expect) were calculated with PacBio's *in silico* kinetic reference computational control model. The accuracy of modification detection using this model was increased by comparing the observed IPD ratios to the expected signatures of the bacterial modification types m6A and m4C. Sequence motif cluster analysis was performed with PacBio Motif finder v1 using the default quality value (QV) cutoff of 30.

Strand specificity and genomic distribution of modified bases were determined with BedOps v2.4.32 (78). To determine localized differences in motif modification, the genome was split into 1,000-bp segments with a 250-bp overlapping sequence. BED files of genomic modifications were intersected with genomic features and/or segments to obtain counting statistics. The metagene analysis of m6A distributions within genes was performed with deepTools2 (79).

RNA isolation and sequencing. Three independent cultures each of the wild-type and MTase mutant strains were grown to mid-late-log phase (3×10^7 cells/ml) for RNA isolation. RNA was isolated by first adding 20 ml RNAprotect (Qiagen) to 10 ml culture to stabilize the transcripts during processing. Cells were collected by centrifugation, resuspended in 1 ml of prewarmed (65°C) TRIzol reagent (Invitrogen), and frozen at −80°C overnight. Cell suspensions were thawed at room temperature, and RNA was isolated with the Direct-zol RNA minikit (Zymo Research) according to the manufacturer's instructions. RNA concentration was determined by a Qubit 2.0 fluorometer (Life Technologies), and RNA integrity was verified by microfluidic-based capillary electrophoresis with an Agilent 2100 Bioanalyzer (RNA integrity number [RIN] $\geq$ 9.8 for all samples).

Directional cDNA libraries were prepared from 5 μg of purified RNA as the input by using the ScriptSeq complete bacteria kit and ScriptSeq index PCR primers (Illumina) according to the manufacturer's instructions. Libraries were analyzed on the Bioanalyzer to ensure appropriate size distributions. The cDNA library from one of the three BbΔe02 strain samples did not pass library quality control (QC) and was omitted. The remaining 8 indexed cDNA libraries were pooled and sequenced on two runs using the Illumina MiSeq with 150-cycle V3 kits (75-bp paired-end [PE] reads were collected). The reads from each sample in the pool were demultiplexed with Illumina CASAVA software v1.8, and fastq files from the two sequencing runs were combined for each sample prior to analysis.

For the analysis of transcriptome sequencing (RNA-seq) data, adapters were removed from the sequencing reads by Trimmomatic (80). The reads were aligned and counted with a transcriptome

reference compiled from the *B. burgdorferi* genome (RefSeq numbers AE000783 to AE000794 and AE001575 to AE001584) as well as with the ncRNAs described by Arnold et al. (45) by using Salmon 0.81 (81). A differential expression analysis was conducted on the raw read counts with DEseq2 (82). Genes in the mutant strains were considered to have significantly different expression compared to that of the wild-type strain at a false discovery rate of ≤0.01. Plasmids lp5, cp9, cp32-6, lp38, and lp56 were removed prior to differential expression analysis, as they are missing in at least one of the strains analyzed.

Data visualization. Data generated from DNA and RNA sequencing analyses were visualized with R v.3.3.0 (https://www.R-project.org/) using the following packages: ggseqlogo (83) for sequence logos, Rcircos (84) for the circos plot, and ggplot2 (85) for bar plots, box plots, and MA plots.

Accession number(s). Sequences have been deposited in the NCBI GEO sequence read archive database under accession number GSE115308.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JB .00395-18.

SUPPLEMENTAL FILE 1, XLSX file, 0.6 MB.
SUPPLEMENTAL FILE 2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Hinckley AF, Connally NP, Meek JI, Johnson BJ, Kemperman MM, Feldman KA, White JL, Mead PS. 2014. Lyme disease testing by large commercial laboratories in the United States. Clin Infect Dis 59:676–681. https://doi.org/10.1093/cid/ciu397.

2. Schwartz AM, Hinckley AF, Mead PS, Hook SA, Kugeler KJ. 2017. Surveillance for Lyme disease—United States, 2008–2015. MMWR Surveill Summ 66:1–12. https://doi.org/10.15585/mmwr.ss6622a1.

3. Bratton RL, Whiteside JW, Hovan MJ, Engle RL, Edwards FD. 2008. Diagnosis and treatment of Lyme disease. Mayo Clin Proc 83:566–571. https://doi.org/10.4065/83.5.566.

4. Rudenko N, Golovchenko M, Grubhoffer L, Oliver JH, Jr. 2011. Updates on *Borrelia burgdorferi* sensu lato complex with respect to public health. Ticks Tick Borne Dis 2:123–128. https://doi.org/10.1016/j.ttbdis.2011.04 .002.

5. Strle K, Drouin EE, Shen S, El Khoury J, McHugh G, Ruzic-Sabljic E, Strle F, Steere AC. 2009. *Borrelia burgdorferi* stimulates macrophages to secrete higher levels of cytokines and chemokines than *Borrelia afzelii* or *Borrelia garinii*. J Infect Dis 200:1936–1943. https://doi.org/10.1086/ 648091.

6. Jones KL, Muellegger RR, Means TK, Lee M, Glickstein LJ, Damle N, Sikand VK, Luster AD, Steere AC. 2008. Higher mRNA levels of chemokines and cytokines associated with macrophage activation in erythema migrans skin lesions in patients from the United States than in patients from Austria with Lyme borreliosis. Clin Infect Dis 46:85–92. https://doi.org/ 10.1086/524022.

7. Carlsson SA, Granlund H, Jansson C, Nyman D, Wahlberg P. 2003. Characteristics of erythema migrans in *Borrelia afzelii* and *Borrelia garinii* infections. Scand J Infect Dis 35:31–33. https://doi.org/10.1080/ 0036554021000026978.

8. Wormser GP, Brisson D, Liveris D, Hanincova K, Sandigursky S, Nowakowski J, Nadelman RB, Ludin S, Schwartz I. 2008. *Borrelia burgdorferi* genotype predicts the capacity for hematogenous dissemination during early Lyme disease. J Infect Dis 198:1358–1364. https://doi.org/10.1086/ 592279.

9. Jones KL, Glickstein LJ, Damle N, Sikand VK, McHugh G, Steere AC. 2006. *Borrelia burgdorferi* genetic markers and disseminated disease in patients with early Lyme disease. J Clin Microbiol 44:4407–4413. https:// doi.org/10.1128/JCM.01077-06.

10. Lagal V, Postic D, Ruzic-Sabljic E, Baranton G. 2003. Genetic diversity among *Borrelia* strains determined by single-strand conformation polymorphism analysis of the ospC gene and its association with invasive-

ness. J Clin Microbiol 41:5059–5065. https://doi.org/10.1128/JCM.41.11 .5059-5065.2003.

11. Jones KL, McHugh GA, Glickstein LJ, Steere AC. 2009. Analysis of *Borrelia burgdorferi* genotypes in patients with Lyme arthritis: high frequency of ribosomal RNA intergenic spacer type 1 strains in antibiotic-refractory arthritis. Arthritis Rheum 60:2174–2182. https://doi.org/10.1002/art.24812.

12. Wang G, Ojaimi C, Iyer R, Saksenberg V, McClain SA, Wormser GP, Schwartz I. 2001. Impact of genotypic variation of *Borrelia burgdorferi* sensu stricto on kinetics of dissemination and severity of disease in C3H/HeJ mice. Infect Immun 69:4303–4312. https://doi.org/10.1128/IAI .69.7.4303-4312.2001.

13. Dykhuizen DE, Brisson D, Sandigursky S, Wormser GP, Nowakowski J, Nadelman RB, Schwartz I. 2008. The propensity of different *Borrelia burgdorferi* sensu stricto genotypes to cause disseminated infections in humans. Am J Trop Med Hyg 78:806–810. https://doi.org/10.4269/ajtmh .2008.78.806.

14. Strle K, Jones KL, Drouin EE, Li X, Steere AC. 2011. *Borrelia burgdorferi* RST1 (OspC type A) genotype is associated with greater inflammation and more severe Lyme disease. Am J Pathol 178:2726–2739. https://doi .org/10.1016/j.ajpath.2011.02.018.

15. Xiang X, Yang Y, Du J, Lin T, Chen T, Yang XF, Lou Y. 2017. Investigation of ospC Expression Variation among *Borrelia burgdorferi* Strains. Front Cell Infect Microbiol 7:131. https://doi.org/10.3389/fcimb.2017.00131.

16. Edwards JR, Yarychkivska O, Boulard M, Bestor TH. 2017. DNA methylation and DNA methyltransferases. Epigenetics Chromatin 10:23. https:// doi.org/10.1186/s13072-017-0130-8.

17. Sánchez-Romero MA, Cota I, Casadesus J. 2015. DNA methylation in bacteria: from the methyl group to the methylome. Curr Opin Microbiol 25:9–16. https://doi.org/10.1016/j.mib.2015.03.004.

18. Loenen WA, Dryden DT, Raleigh EA, Wilson GG, Murray NE. 2014. Highlights of the DNA cutters: a short history of the restriction enzymes. Nucleic Acids Res 42:3–19. https://doi.org/10.1093/nar/gkt990.

19. Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, Froula J, Kang DD, Malmstrom RR, Morgan RD, Posfai J, Singh K, Visel A, Wetmore K, Zhao Z, Rubin EM, Korlach J, Pennacchio LA, Roberts RJ. 2016. The epigenomic landscape of prokaryotes. PLoS Genet 12: e1005854. https://doi.org/10.1371/journal.pgen.1005854.

20. Pingoud A, Wilson GG, Wende W. 2014. Type II restriction endonucleases–a historical perspective and more. Nucleic Acids Res 42:7489–7527. https:// doi.org/10.1093/nar/gku447.

21. Seshasayee AS, Singh P, Krishna S. 2012. Context-dependent conserva-

tion of DNA methyltransferases in bacteria. Nucleic Acids Res 40: 7066–7073. https://doi.org/10.1093/nar/gks390.

22. Vasu K, Nagaraja V. 2013. Diverse functions of restriction-modification systems in addition to cellular defense. Microbiol Mol Biol Rev 77:53–72. https://doi.org/10.1128/MMBR.00044-12.

23. Korlach J, Turner SW. 2012. Going beyond five bases in DNA sequencing. Curr Opin Struct Biol 22:251–261. https://doi.org/10.1016/j.sbi.2012.04 .002.

24. Lawrenz MB, Kawabata H, Purser JE, Norris SJ. 2002. Decreased electroporation efficiency in Borrelia burgdorferi containing linear plasmids lp25 and lp56: impact on transformation of infectious B. burgdorferi. Infect Immun 70:4798–4804. https://doi.org/10.1128/IAI.70.9.4798-4804.2002.

25. Rego RO, Bestor 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.

26. 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.

27. Stewart PE, Hoff J, Fischer E, Krum JG, Rosa PA. 2004. Genome-wide transposon mutagenesis of Borrelia burgdorferi for identification of phenotypic mutants. Appl Environ Microbiol 70:5973–5979. https://doi.org/ 10.1128/AEM.70.10.5973-5979.2004.

28. Botkin DJ, Abbott AN, Stewart PE, Rosa PA, Kawabata H, Watanabe H, Norris SJ. 2006. Identification of potential virulence determinants by Himar1 transposition of infectious Borrelia burgdorferi B31. Infect Immun 74:6690–6699. https://doi.org/10.1128/IAI.00993-06.

29. Karna SL, Sanjuan E, Esteve-Gassent MD, Miller CL, Maruskova M, Seshu J. 2011. CsrA modulates levels of lipoproteins and key regulators of gene expression critical for pathogenic mechanisms of Borrelia burgdorferi. Infect Immun 79:732–744. https://doi.org/10.1128/IAI.00882-10.

30. Hyde JA, Weening EH, Chang M, Trzeciakowski JP, Hook M, Cirillo JD, Skare JT. 2011. Bioluminescent imaging of Borrelia burgdorferi in vivo demonstrates that the fibronectin-binding protein BBK32 is required for optimal infectivity. Mol Microbiol 82:99–113. https://doi.org/10.1111/j .1365-2958.2011.07801.x.

31. Ellis TC, Jain S, Linowski AK, Rike K, Bestor A, Rosa PA, Halpern M, Kurhanewicz S, Jewett MW. 2014. Correction. In vivo expression technology identifies a novel virulence factor critical for Borrelia burgdorferi persistence in mice. PLoS Pathog 10:e1004260. https://doi.org/10.1371/ journal.ppat.1004260.

32. Casselli T, Bankhead T. 2015. Use of in vivo expression technology for the identification of putative host adaptation factors of the Lyme disease spirochete. J Mol Microbiol Biotechnol 25:349–361. https://doi.org/10 .1159/000439305.

33. Shi Y, Xu Q, Seemanapalli SV, McShan K, Liang FT. 2006. The dbpBA locus of Borrelia burgdorferi is not essential for infection of mice. Infect Immun 74:6509–6512. https://doi.org/10.1128/IAI.00740-06.

34. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, et al. 2009. Real-time DNA sequencing from single polymerase molecules. Science 323:133–138. https://doi.org/10.1126/science.1162986.

35. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, 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.

36. Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts RJ, Korlach J. 2012. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. Nucleic Acids Res 40:e29. https://doi.org/10.1093/nar/gkr1146.

37. Roberts RJ, Vincze T, Posfai J, Macelis D. 2015. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res 43:D298–D299. https://doi.org/10.1093/nar/gku1046.

38. Morgan RD, Dwinell EA, Bhatia TK, Lang EM, Luyten YA. 2009. The MmeI family: type II restriction-modification enzymes that employ singlestrand modification for host protection. Nucleic Acids Res 37:5208–5221. https://doi.org/10.1093/nar/gkp534.

39. Morgan RD, Bhatia TK, Lovasco L, Davis TB. 2008. MmeI: a minimal type

II restriction-modification system that only modifies one DNA strand for host protection. Nucleic Acids Res 36:6558–6570. https://doi.org/10 .1093/nar/gkn711.

40. Picardeau M, Lobry JR, Hinnebusch BJ. 2000. Analyzing DNA strand compositional asymmetry to identify candidate replication origins of Borrelia burgdorferi linear and circular plasmids. Genome Res 10: 1594–1604. https://doi.org/10.1101/gr.124000.

41. Casjens S, Palmer N, Van Vugt R, Huang WH, Stevenson B, Rosa P, Lathigra R, Sutton G, Peterson J, Dodson RJ, Haft D, Hickey E, Gwinn M, White O, Fraser CM. 2000. A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi. Mol Microbiol 35:490–516. https://doi.org/10.1046/j.1365-2958.2000.01698.x.

42. Eggers CH, Samuels DS. 1999. Molecular evidence for a new bacteriophage of Borrelia burgdorferi. J Bacteriol 181:7308–7313.

43. Eggers CH, Casjens S, Hayes SF, Garon CF, Damman CJ, Oliver DB, Samuels DS. 2000. Bacteriophages of spirochetes. J Mol Microbiol Biotechnol 2:365–373.

44. Zhang H, Marconi RT. 2005. Demonstration of cotranscription and 1-methyl-3-nitroso-nitroguanidine induction of a 30-gene operon of Borrelia burgdorferi: evidence that the 32-kilobase circular plasmids are prophages. J Bacteriol 187:7985–7995. https://doi.org/10.1128/JB.187.23 .7985-7995.2005.

45. Arnold WK, Savage CR, Brissette CA, Seshu J, Livny J, Stevenson B. 2016. RNA-Seq of Borrelia burgdorferi in multiple phases of growth reveals insights into the dynamics of gene expression, transcriptome architecture, and noncoding RNAs. PLoS One 11:e0164165. https://doi.org/10 .1371/journal.pone.0164165.

46. Wang W, Xu L, Hu L, Chong J, He C, Wang D. 2017. Epigenetic DNA modification N(6)-methyladenine causes site-specific RNA polymerase II transcriptional pausing. J Am Chem Soc 139:14436–14442. https://doi .org/10.1021/jacs.7b06381.

47. Xu L, Wang W, Chong J, Shin JH, Xu J, Wang D. 2015. RNA polymerase II transcriptional fidelity control and its functional interplay with DNA modifications. Crit Rev Biochem Mol Biol 50:503–519. https://doi.org/10 .3109/10409238.2015.1087960.

48. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK, Gwinn M, Dougherty B, Tomb JF, Fleischmann RD, Richardson D, Peterson J, Kerlavage AR, Quackenbush J, Salzberg S, Hanson M, van Vugt R, Palmer N, Adams MD, Gocayne J, Weidman J, Utterback T, Watthey L, McDonald L, Artiach P, Bowman C, Garland S, Fujii C, Cotton MD, Horst K, Roberts K, Hatch B, Smith HO, Venter JC. 1997. Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. Nature 390:580–586. https://doi.org/10.1038/37551.

49. 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.

50. James AE, Rogovskyy AS, Crowley MA, Bankhead T. 2016. Characterization of a DNA adenine methyltransferase gene of Borrelia hermsii and its dispensability for murine infection and persistence. PLoS One 11: e0155798. https://doi.org/10.1371/journal.pone.0155798.

51. Murray IA, Clark TA, Morgan RD, Boitano M, Anton BP, Luong K, Fomenkov A, Turner SW, Korlach J, Roberts RJ. 2012. The methylomes of six bacteria. Nucleic Acids Res 40:11450–11462. https://doi.org/10.1093/ nar/gks891.

52. Pingoud A, Fuxreiter M, Pingoud V, Wende W. 2005. Type II restriction endonucleases: structure and mechanism. Cell Mol Life Sci 62:685–707. https://doi.org/10.1007/s00018-004-4513-1.

53. Fang G, Munera D, Friedman DI, Mandlik A, Chao MC, Banerjee O, Feng Z, Losic B, Mahajan MC, Jabado OJ, Deikus G, Clark TA, Luong K, Murray IA, Davis BM, Keren-Paz A, Chess A, Roberts RJ, Korlach J, Turner SW, Kumar V, Waldor MK, Schadt EE. 2012. Genome-wide mapping of methylated adenine residues in pathogenic Escherichia coli using singlemolecule real-time sequencing. Nat Biotechnol 30:1232–1239. https:// doi.org/10.1038/nbt.2432.

54. Anjum A, Brathwaite KJ, Aidley J, Connerton PL, Cummings NJ, Parkhill J, Connerton I, Bayliss CD. 2016. Phase variation of a type IIG restrictionmodification enzyme alters site-specific methylation patterns and gene expression in Campylobacter jejuni strain NCTC11168. Nucleic Acids Res 44:4581–4594. https://doi.org/10.1093/nar/gkw019.

55. Miller CB, Pierle SA, Brayton KA, Ochoa JN, Shah DH, Lahmers KK. 2014. Transcriptional profiling of a cross-protective Salmonella enterica serovar Typhimurium UK-1 dam mutant identifies a set of genes more transcriptionally active compared to wild-type, and stably transcribed across

biologically relevant microenvironments. Pathogens 3:417–436. https://doi.org/10.3390/pathogens3020417.

56. Radolf JD, Caimano MJ, Stevenson B, Hu LT. 2012. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. Nat Rev Microbiol 10:87–99. https://doi.org/10.1038/nrmicro2714.

57. Samuels DS. 2011. Gene regulation in *Borrelia burgdorferi*. Annu Rev Microbiol 65:479–499. https://doi.org/10.1146/annurev.micro.112408.134040.

58. Stevenson B, Seshu J. 2018. Regulation of gene and protein expression in the Lyme disease spirochete. Curr Top Microbiol Immunol 415:83–112. https://doi.org/10.1007/82_2017_49.

59. Lybecker MC, Samuels DS. 2007. Temperature-induced regulation of RpoS by a small RNA in *Borrelia burgdorferi*. Mol Microbiol 64:1075–1089. https://doi.org/10.1111/j.1365-2958.2007.05716.x.

60. Dulebohn DP, Hayes BM, Rosa PA. 2014. Global repression of host-associated genes of the Lyme disease spirochete through post-transcriptional modulation of the alternative sigma factor RpoS. PLoS One 9:e93141. https://doi.org/10.1371/journal.pone.0093141.

61. Caimano MJ, Iyer R, Eggers CH, Gonzalez C, Morton EA, Gilbert MA, Schwartz I, Radolf JD. 2007. Analysis of the RpoS regulon 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.

62. Xu Q, McShan K, Liang FT. 2007. Identification of an ospC operator critical for immune evasion of *Borrelia burgdorferi*. Mol Microbiol 64:220–231. https://doi.org/10.1111/j.1365-2958.2007.05636.x.

63. Arnold WK, Savage CR, Lethbridge KG, Smith TC, II, Brissette CA, Seshu J, Stevenson B. 2018. Transcriptomic insights on the virulence-controlling CsrA, BadR, RpoN, and RpoS regulatory networks in the Lyme disease spirochete. PLoS One 13:e0203286. https://doi.org/10.1371/journal.pone.0203286.

64. Ryjenkov DA, Tarutina M, Moskvin OV, Gomelsky M. 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. J Bacteriol 187:1792–1798. https://doi.org/10.1128/JB.187.5.1792-1798.2005.

65. Sultan SZ, Pitzer JE, Boquoi T, Hobbs G, Miller MR, Motaleb MA. 2011. Analysis of the HD-GYP domain cyclic dimeric GMP phosphodiesterase reveals a role in motility and the enzootic life cycle of *Borrelia burgdorferi*. Infect Immun 79:3273–3283. https://doi.org/10.1128/IAI.05153-11.

66. He M, Ouyang Z, Troxell B, Xu H, Moh A, Piesman J, Norgard MV, Gomelsky M, Yang XF. 2011. Cyclic di-GMP is essential for the survival of the Lyme disease spirochete in ticks. PLoS Pathog 7:e1002133. https://doi.org/10.1371/journal.ppat.1002133.

67. Caimano MJ, Dunham-Ems S, Allard AM, Cassera MB, Kenedy M, Radolf JD. 2015. Cyclic di-GMP modulates gene expression in Lyme disease spirochetes at the tick-mammal interface to promote spirochete survival during the blood meal and tick-to-mammal transmission. Infect Immun 83:3043–3060. https://doi.org/10.1128/IAI.00315-15.

68. Adams PP, Flores Avile C, Popitsch N, Bilusic I, Schroeder R, Lybecker M, Jewett MW. 2017. *In vivo* expression technology and 5′ end mapping of the *Borrelia burgdorferi* transcriptome identify novel RNAs expressed during mammalian infection. Nucleic Acids Res 45:775–792. https://doi.org/10.1093/nar/gkw1180.

69. Jacobs MB, Norris SJ, Phillippi-Falkenstein KM, Philipp MT. 2006. Infectivity of the highly transformable BBE02- lp56- mutant of *Borrelia burgdorferi*, the Lyme disease spirochete, via ticks. Infect Immun 74:3678–3681. https://doi.org/10.1128/IAI.00043-06.

70. Revel AT, Talaat AM, Norgard MV. 2002. DNA microarray analysis of differential gene expression in *Borrelia burgdorferi*, the Lyme disease spirochete. Proc Natl Acad Sci U S A 99:1562–1567. https://doi.org/10.1073/pnas.032667699.

71. Balbontín R, Rowley G, Pucciarelli MG, Lopez-Garrido J, Wormstone Y, Lucchini S, Garcia-Del Portillo F, Hinton JC, Casadesus J. 2006. DNA adenine methylation regulates virulence gene expression in *Salmonella enterica* serovar Typhimurium. J Bacteriol 188:8160–8168. https://doi.org/10.1128/JB.00847-06.

72. Marinus MG, Casadesus J. 2009. Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS Microbiol Rev 33:488–503. https://doi.org/10.1111/j.1574-6976.2008.00159.x.

73. Elias AF, Stewart PE, Grimm D, Caimano MJ, Eggers CH, Tilly K, Bono JL, Akins DR, Radolf JD, Schwan TG, Rosa P. 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.

74. Barbour AG. 1984. Isolation and cultivation of Lyme disease spirochetes. Yale J Biol Med 57:521–525.

75. Samuels DS. 1995. Electrotransformation of the spirochete *Borrelia burgdorferi*. Methods Mol Biol 47:253–259.

76. Bono JL, Elias AF, Kupko JJ, III, Stevenson B, Tilly K, Rosa P. 2000. Efficient targeted mutagenesis in *Borrelia burgdorferi*. J Bacteriol 182:2445–2452. https://doi.org/10.1128/JB.182.9.2445-2452.2000.

77. Bunikis I, Kutschan-Bunikis S, Bonde M, Bergstrom S. 2011. Multiplex PCR as a tool for validating plasmid content of *Borrelia burgdorferi*. J Microbiol Methods 86:243–247. https://doi.org/10.1016/j.mimet.2011.05.004.

78. Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, Sandstrom R, Humbert R, Stamatoyannopoulos JA. 2012. BEDOPS: high-performance genomic feature operations. Bioinformatics 28:1919–1920. https://doi.org/10.1093/bioinformatics/bts277.

79. Ramírez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44:W160–W165. https://doi.org/10.1093/nar/gkw257.

80. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

81. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417–419. https://doi.org/10.1038/nmeth.4197.

82. 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.

83. Wagih O. 2017. ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics 33:3645–3647. https://doi.org/10.1093/bioinformatics/btx469.

84. Zhang H, Meltzer P, Davis S. 2013. RCircos: an R package for Circos 2D track plots. BMC Bioinformatics 14:244. https://doi.org/10.1186/1471-2105-14-244.

85. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY. http://ggplot2.org.