Characterization of Undermethylated Sites in *Vibrio cholerae*

Ankur B. Dalia, David W. Lazinski, Andrew Camilli

Howard Hughes Medical Institute and Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts, USA

The activities of DNA methyltransferases are important for a variety of cellular functions in bacteria. In this study, we developed a modified high-throughput technique called methyl homopolymer tail mediated sequencing (methyl HTM-seq) to identify the undermethylated sites in the *Vibrio cholerae* genome for the two DNA methyltransferases, Dam, an adenine methyltransferase, and VchM, a cytosine methyltransferase, during growth in rich medium *in vitro*. Many of the undermethylated sites occurred in intergenic regions, and for most of these sites, we identified the transcription factors responsible for undermethylation. This confirmed the presence of previously hypothesized DNA-protein interactions for these transcription factors and provided insight into the biological state of these cells during growth *in vitro*. DNA adenine methylation has previously been shown to mediate heritable epigenetic switches in gene regulation. However, none of the undermethylated Dam sites tested showed evidence of regulation by this mechanism. This study is the first to identify undermethylated adenines and cytosines genomewide in a bacterium using second-generation sequencing technology.

To gain further insight into the role of methylation in *V. cholerae*, we sought to define the undermethylated sites for both Dam and VchM in the O1 serogroup, El Tor biotype clinical isolate E7946. Since undermethylation can be caused by transcription factor binding, this analysis would potentially define novel transcription factor binding sites that overlap Dam and VchM methylation sites in the genome. Additionally, since methyladenine can regulate gene expression, this analysis may also identify regions in the genome that are regulated by Dam methylation.

**MATERIALS AND METHODS**

**Bacterial strains and culture conditions.** The *V. cholerae* strain used in this study was a streptomycin-resistant variant of the clinical isolate E7946. Cells were usually grown at 37°C in Luria Bertani (LB) broth and on LB agar. Where indicated, cells were also grown at 37°C in M9 minimal medium supplemented with a carbon source as indicated below at a final concentration of 0.4%. When appropriate, cultures were supplemented with streptomycin (100 μg/ml) and/or kanamycin (100 μg/ml) as appropriate. To study lacZ fusions on solid medium, plates were supplemented with 40 μg/ml 5-bromo-4-chloro-3-indolyl-beta-D-galactoside (X-Gal). The PVCA0063 lacZ fusion strain produced dark-blue colonies on LB plates, so in order to reduce the amount of product to enable increases in LacZ expression to be observed, 0.3 mM PETG (phenyl-ethyl-beta-D-thiogalactoside, a competitive inhibitor of X-Gal) was added to the plates, which was the minimum amount required to see light-blue colonies. This lacZ fusion strain was light blue on LB plates containing 200 μM FeSO₄, so PETG was not added to this medium.

**Methyl homopolymer tail mediated-seq (methyl HTM-seq).** Genomic DNA (gDNA) was isolated using the DNeasy blood and tissue kit (QiaGen) according to the manufacturer’s instructions. After isolation, between 1 and 2 μg of gDNA was sheared to ~400 bp (range 200 to 600 bp) using a prechilled Branson high-intensity cup horn sonifier (Branson)
for 2 min at 50% intensity with a 5-s-on/5-s-off duty cycle. Homopolymer tails of cytosine were added to the 3’ ends of all sheared molecules using terminal deoxynucleotidyl transferase (TdT) according to the manufacturer’s instructions (Promega). For these reactions, a 20:1 mixture of dCTP and ddCTP (a chain terminator) was used to generate C-tails of approximately 20 cytosines. After C-tailing, reaction mixtures were run through a Performa spin column (Edge Biosystems) according to the manufacturer’s instructions to remove excess nucleotides and to desalt reaction mixtures. The eluate was digested using the appropriate methyl-sensitivity restriction enzyme (MeSR). For Dam, the MeSR used was MboI, while for VchM, the MeSR used was BsrFI (New England BioLabs). After digestion, samples were heat inactivated when appropriate and run through a Performa spin column to desalt reactions. The eluate was then ligated to the tIL1 adaptor (final concentration of 1 μM) using the Quick Ligase kit (New England BioLabs), in a ligation mixture volume of 60 μl, according to the manufacturer’s instructions. The tIL1 adaptor was generated by annealing the oligonucleotides ABD013 and ABD013D for Dam and ABD013 and ABD013B for VchM (see Table S1 in the supplemental material). Samples were run through a Performa spin column to desalt reaction mixtures. Then, 6 μl of the eluate was used as the template in PCRs to amplify the samples, using OLJ 131 as the forward primer and either OLJ 573 (Dam) or BC33G (VchM) as the reverse primer (see Table 1 in the supplemental material). These forward and reverse primers contain the sequences specific for capture and sequencing on the Illumina HiSeq2000 platform (Illumina). Also, the reverse primers used provide a unique barcode index sequence that can be used to multiplex samples onto a single lane in the Illumina flow cell. After PCR, the DNA concentrations of samples were determined on a Nanodrop 2000 spectrophotometer (Nanodrop) and submitted for sequencing at the Tufts University Core Facility via single-end 50-bp reads on the Illumina HiSeq2000.

After sequencing and demultiplexing, reads were further filtered to identify sequences that represent true Dam and VchM sites. For Dam data, reads were filtered to obtain sequences that started with GATC, while for the VchM data, reads were filtered to obtain sequences that started with either CCGG or CCGGT. The reads were then trimmed to a length of 21 bp and mapped to the O1 El Tor N16961 genome containing the sequence for the K139 prophage inserted into chromosome 1 (since this prophage sequence is absent from N16961 but present in the E7946 strain used in this study [8]), using the program Bowtie, and no mismatches were allowed during mapping (9). Finally, the total numbers of forward and reverse reads mapping to methylation sites were determined.

**MeSR digestion and qPCR for characterization of undermethylated sites**. Assays were performed essentially as previously described (10). Briefly, between 10 and 100 ng of gDNA was digested using a MeSR in a final reaction mixture volume of 20 μl. For Dam, the MeSR used was MboI, while for VchM, the MeSR used was Mspl. After digestion, reaction mixtures were heat inactivated when appropriate. Then, 2 μl of this digest (1 to 10 ng) was used as the template for quantitative PCR (qPCR) using primers that span a methylation site of interest. All primers used for MeSR digestion and qPCR are listed in Table S1 in the supplemental material. An uncultured control reaction mixture (no-enzyme control) was run for every sample. The abundance of DNA in all samples was determined relative to a standard curve generated using dilutions of purified gDNA. Reaction mixtures were run on an Mv300SP qPCR instrument (Stratagene) using the dye incorporation method (SYBR green) and analyzed using MxPro qPCR software (Stratagene).

**Generation of mutant strains**. Deletion mutants were generated using the natural competence of *V. cholerae*, and PCR products generated by splicing overlap extension (SOE) PCR essentially as previously described (11, 12). The oligonucleotides used for SOE PCRs are shown in Table S1 in the supplemental material. For all SOE products, PCRs were performed to generate an up arm (primers F1/R1), a down arm (primers F2/R2), and a kanamycin resistance cassette (primers FRT-Kan-F/FRT-Kan-R). The PCR products for the up arm, down arm, and kanamycin resistance cassette were then mixed 1:1:1 (~50 ng each) and used as the template in a PCR mixture using the F1 and R2 primers to generate the full-length SOE product. This product was then purified using QIAquick PCR cleanup columns (Qiagen), and 1- to 5-μg amounts were used to transform *V. cholerae* via natural competence as previously described (11). The lacZ fusion strain was generated by a similar method. A SOE product was generated to delete the endogenous lacZ (VC2338) as described above, and a second SOE product was generated to transcriptionally fuse the VCA0063 promoter to the *V. cholerae* lacZ gene at the VCA0063 locus. Both constructs were simultaneously transformed into *V. cholerae* via natural transformation. All mutant strains were confirmed by PCR and/or sequencing.

**Purification of Fur and electrophoretic mobility shift assays (EMSA)**. The gene encoding the *V. cholerae* Fur protein was amplified using primers ABD084 and ABD085 and cloned into the Ndel and BamHI sites of pET15b-His-Tev to generate a Fur protein containing a His tag that can be removed by cleavage with the Tev protease as previously described (see Table S1 in the supplemental material) (13). This expression vector was transformed into *E. coli* BL21 (DE3), a single colony was picked, and the sequence of the insertion was confirmed by PCR and sequencing. To induce expression, cells were grown to an optical density at 600 nm (OD600) of ~0.8 in LB medium containing 100 μg/ml ampicillin at 37°C with shaking (250 rpm), and then isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the culture to a final concentration of 1 mM and the culture incubated for 16 h at 18°C shaking (250 rpm). Protein was then purified from these cells essentially as previously described using a Ni-nitrilotriacetic acid (NTA) column (13). The His tag was then cleaved from Fur by incubating purified protein with Tev protease at 4°C overnight as previously described (13). Untagged Fur protein was then stored in 10 mM Tris, pH 8.0, 250 mM NaCl, 5 mM β-mercaptoethanol, and 50% glycerol at −20°C for short-term storage (up to 1 week) or at −80°C for long-term storage (tested up to 2 months).

The primers used to generate probes are listed in Table S1 in the supplemental material. For the probe of *VCA0063*, oligonucleotides end labeled with Cy5 were ordered (IDT) and were annealed in vitro, while for *pdbA* and *psphA*, Cy5 was incorporated into probes during PCR using Cy5-labeled dCTP (GE Healthcare). For the latter, we specifically determined the titer of the ratio of unlabeled dCTP to Cy5-labeled dCTP added to PCR products in the incorporation of only 1 or 2 labeled C residues, as we found this to be sufficient to result in a robustly labeled EMSA probe. Where indicated, EMSA probes were in vitro methylated with Dam methylease according to the manufacturer’s instruction (NEB). After methylation, reaction mixtures were run through a Performa spin column to remove excess 5-adenosylmethionine and to desalt the mixtures.

The binding reaction was done in a 20-μl mixture that contained probe (2 nM), Fur protein (as indicated below), 10 mM Tris-borate buffer (5.4 g/liter Tris base and 80 mM NaCl), 1 mM MgCl₂, 100 μM MnCl₂, 2 mM dithiothreitol (DTT), 100 μg/ml bovine serum albumin (BSA), 5 μg sheared calf thymus DNA, and 10% glycerol. The reaction mixtures were incubated for 30 min at room temperature prior to loading onto prerun 6% native polyacrylamide gels. Gels were made in 0.5% Tris-borate buffer (5.4 g/liter Tris base and 2.25 g/liter boric acid) and prerun for 1 h in 0.5% Tris-borate buffer containing 100 μM MnCl₂. Gels were imaged using the Cy5 setting on the FLA-9000IR instrument (GE Healthcare Life Sciences).

**Transcript abundance by qRT-PCR**. Total cellular RNA was purified from *V. cholerae* using the RNEasy minikit (Qiagen) and then treated with DNase from the TURBO DNA-free kit (Ambion) according to the manufacturer’s instructions. Next, cDNA was generated from 1 μg of treated RNA using random priming with the iScript cDNA synthesis kit in a volume of 20 μl according to the manufacturer’s instructions (Bio-Rad). Two microliters of each sample was loaded in quadruplicate onto a qPCR plate. Two wells were used to assess the transcript abundance of the query ampiclon (genes VCA0063 or VC1784), while the other two wells were used to assess the transcript abundance of the housekeeping gene *rpoB*. The primers used for all quantitative reverse transcriptase PCR (qRT-PCR) experiments are listed in Table S1 in the supplemental material, and...
qPCR was performed as described above. The relative abundance of transcripts was determined using the standard curve method: for each reaction plate, two standard curves of gDNA were included and tested with the amplicon-specific primers and the primers for *rpoB*. The expression of each amplicon was determined relative to the expression of *rpoB*. To compare between different growth conditions, expression is shown relative to one of the conditions (usually relative to growth in LB). For all samples, a control reaction mixture lacking reverse transcriptase was run to confirm that RNA samples were not contaminated with gDNA.

**Statistical comparisons and DNA analysis.** All data were plotted using GraphPad Prism version 5.0 (GraphPad), and statistical comparisons were made as indicated in figure legends. Analysis of DNA was performed using CLC Main Workbench 6 (CLC Bio).

**RESULTS**

**Identification of undermethylated sites in the *V. cholerae* genome by a modified high-throughput approach.** To identify the undermethylated sites in both chromosomes of the *V. cholerae* O1 El Tor strain E7946, we undertook a high-throughput approach using methylation-sensitive restriction enzymes (MeSR) and next-generation sequencing (NGS). This approach is similar to the methyl-seq technique used to study methylation profiles in eukaryotes (14). Methyl-seq uses MeSRs to cut undermethylated sites in the genome and then ligates the ends generated to adaptors that allow for NGS. Thus, a single unmethylated site should be represented by two independent sequences (reads). The methyl-seq strategy is limited, however, since areas where unmethylated sites are rare and thus widely spaced in the genome will not be represented in the output due to the constraints of size selection for sequencing on NGS platforms like Illumina (14, 15). This can result in these sites being completely absent in the output or can result in the loss of one of the two paired sequences that define an unmethylated site (15). To circumvent this problem, we utilized a modification of the recently described homopolymer tail-mediated ligation PCR (HTML-PCR) method for generating Illumina genomic libraries to probe for unmethylated sites in the genome (16). In this approach, DNA is first sheared to ~200 to 600 bp, and then homopolymer C-tails are added to the 3’ ends of all molecules using terminal deoxynucleotidyl transferase (TdT) (Fig. 1A). Next, a MeSR is used to cut all unmethylated sites, and the cut ends are ligated to an adaptor (tIL1) (Fig. 1A). Once ligated, libraries are amplified using oligonucleotide primers containing the sequences necessary for sequencing on the Illumina platform (IL-1 and IL-2) (Fig. 1A). Therefore, we are calling this modified technique methyl homopolymer tail mediated sequencing (methyl HTM-seq). One potential issue that can arise during methyl HTM-seq is that unmethylated sites that are in close proximity to one another (~100 bp apart) may lose one of the sequences that forms a pair for each unmethylated site, since the likelihood of shearing the genomic DNA (gDNA) between these two sites is relatively low. The outside reads that span the two pairs, however, should still be represented and constitute a pseudopair that can be used to infer the unmethylated status of closely spaced methylation sites.

To assess undermethylation of Dam and VchM sites in *V. cholerae*, we used the MeSRs MboI and BsrFI, respectively, which per-
Undermethylated intergenic sites identified by methyl HTM-seq

<table>
<thead>
<tr>
<th>Chromosome Position</th>
<th>Intergenic locus</th>
<th>Frequency</th>
<th>Fold overrepresentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dam data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chrI 1306070</td>
<td>VC1231-VC1232</td>
<td>6.96E-04</td>
<td>26.2</td>
</tr>
<tr>
<td>chrI 1306010</td>
<td>VC1231-VC1232</td>
<td>5.74E-04</td>
<td>21.6</td>
</tr>
<tr>
<td>chrI 1933145</td>
<td>VC1783-VC1784</td>
<td>3.28E-04</td>
<td>12.4</td>
</tr>
<tr>
<td>chrI 1933148</td>
<td>VC1783-VC1784</td>
<td>4.60E-04</td>
<td>17.3</td>
</tr>
<tr>
<td>chrI 2364224</td>
<td>K139p05-K139p04</td>
<td>2.02E-04</td>
<td>7.6</td>
</tr>
<tr>
<td>chrI 2364255</td>
<td>K139p05-K139p04</td>
<td>1.22E-04</td>
<td>4.6</td>
</tr>
<tr>
<td>VchM data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chrI 68810</td>
<td>VCA0602-VCA0603</td>
<td>3.11E-04</td>
<td>11.7</td>
</tr>
<tr>
<td>chrI 68813</td>
<td>VCA0602-VCA0603</td>
<td>1.07E-03</td>
<td>40.3</td>
</tr>
</tbody>
</table>

Undermethylated intergenic sites identified by methyl HTM-seq

Table 1

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position</th>
<th>Intergenic locus</th>
<th>Frequency</th>
<th>Fold overrepresentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>chrI 1306070</td>
<td>VC1231-VC1232</td>
<td>6.96E-04</td>
<td>26.2</td>
<td></td>
</tr>
<tr>
<td>chrI 1306010</td>
<td>VC1231-VC1232</td>
<td>5.74E-04</td>
<td>21.6</td>
<td></td>
</tr>
<tr>
<td>chrI 1933145</td>
<td>VC1783-VC1784</td>
<td>3.28E-04</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>chrI 1933148</td>
<td>VC1783-VC1784</td>
<td>4.60E-04</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>chrI 2364224</td>
<td>K139p05-K139p04</td>
<td>2.02E-04</td>
<td>7.6</td>
<td></td>
</tr>
<tr>
<td>chrI 2364255</td>
<td>K139p05-K139p04</td>
<td>1.22E-04</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>chrI 68810</td>
<td>VCA0602-VCA0603</td>
<td>3.11E-04</td>
<td>11.7</td>
<td></td>
</tr>
<tr>
<td>chrI 68813</td>
<td>VCA0602-VCA0603</td>
<td>1.07E-03</td>
<td>40.3</td>
<td></td>
</tr>
</tbody>
</table>

The position and locus are based on the annotated N16961 genome containing the K139 prophage genome in chromosome I. The unmethylated site is located in the intergenic region between the indicated loci.

Indicates the degree of overrepresentation of the reads at the indicated site relative to the total number of reads mapped.

Indicates the frequency of reads mapping to the indicated site relative to the total number of reads mapped.

The relative abundance of DNA in experimental samples is then determined against a standard curve, and the percentage of undermethylation is determined relative to the undigested control. To characterize the undermethylation of Dam sites, we used the MeSR MboI and the MeDR DpnI, and to characterize the undermethylation of VchM sites, we used the MeSR MspI.

Confirmation of undermethylated sites by MeSR digestion and qPCR. To validate and further characterize undermethylated sites, we used a previously described technique that combines MeSR digestion and qPCR (10). The gDNA is first digested with a MeSR and, optionally, also with a methyltransferase dependent restriction enzyme (MeDR). Digestion at a site reduces the number of intact template molecules spanning the region. Samples are heat inactivated and then used as the template in a qPCR using oligo-nucleotide primers that span the methylation site of interest. The

Identification of transcription factors responsible for undermethylation in V. cholerae. For all undermethylated intergenic sites, we attempted to identify a transcription factor that caused the observed methylation pattern. For the K139 prophage, we found two undermethylated sites between the outward-facing genes encoding the cl repressor and Cox (Fig. 2). In the homologous bacteriophage HP1, the cl repressor is responsible for repressing P\textsubscript{PR}, which is the promoter that drives the expression of cI and the downstream genes involved in lytic growth, while Cox is involved in repressing P\textsubscript{C}, which is the promoter that drives the expression of cl and the downstream genes involved in maintenance of lysogeny (18). Since these sites are undermethylated in E7946, a K139 lysogen, we predict that the repressor responsible for the observed methylation profile is the cl repressor; however, this is difficult to confirm experimentally since a mutant strain lacking this repressor would be nonviable. To determine whether we could identify putative operator sites for the cl repressor based on the location of the undermethylated sites in the K139 prophage, we visually inspected the genomic region between P\textsubscript{PR} and P\textsubscript{C} for palindromic sequences or regions of dyad symmetry. This analysis identified two 10-bp sites that overlap or are in close proximity to the two undermethylated Dam sites and, thus, may constitute operator sites for the cl repressor (Fig. 2). The putative cl operator site located between the two undermethylated Dam sites (tg-I and tg-II) is a perfect palindrome with the sequence GTCA AATTGAC, while the site located upstream from P\textsubscript{PR} deviates from this palindrome at 2 sites (GTCAAT GTCG) (Fig. 2).

Based on the literature and the promoter prediction software PROM, there are two potential repressors that may promote undermethylation of P\textsubscript{VC1784}, these are RpiR and Lrp (19). To test this hypothesis, we generated mutant strains and found that RpiR is fully responsible for undermethylation at the Dam site in this...
It was previously shown that VCA0063 is repressed by Fur, and we have determined by using a \(/H\) mutation that this repressor is responsible for undermethylation of the Dam site in this promoter (Fig. 4) (20). The three undermethylated VchM sites are located within the intergenic space between divergent genes VC0286-VC0287, VC1280-VC1281, and VC1558-VC1559, respectively, and each has a gene encoding a putative repressor protein in close proximity in the genome. Analysis of mutant strains lacking these repressors confirmed that they were fully responsible for undermethylation at these VchM sites (Fig. 4).

Overlapping Fur binding and Dam methylation do not mediate an epigenetic switch. The promoter upstream from VCA0063 controls the expression of ptrB and hutR, two genes involved in heme utilization in \(V.\) cholerae (20). This region contains an undermethylated Dam site that overlaps with the Fur binding site and the \(/H\) consensus sequence of PVCA0063 (Fig. 5A). An almost identical regulatory arrangement is seen in the promoter of the \(scil\) gene cluster, which controls the expression of the type VI secretion system in \(E.\) coli. Recently, it was shown for the \(scil\) promoter that Fur binding results in undermethylation of the Dam site that overlaps the Fur consensus sequence and that Dam methylation reduced the affinity for Fur binding at this site (21). Analogous regulatory arrangements have been shown to promote a heritable epigenetic switch in gene regulation in other systems (2, 22). The best-described example is the control of Pap pilus expression in uropathogenic \(E.\) coli. In this system, the global response regulator Lrp acts as a repressor and competes with Dam to prevent methylation at a Dam site proximal to the promoter driving expression of the pilus. At a low, stochastic frequency within a population of cells, Dam will gain entry to this site and methylate it. This, in turn, inhibits Lrp from binding to the region proximal to the Pap promoter and allows for expression of the pilus. This phenotype is heritable because, upon cell division, DNA becomes hemimethylated, which still results in poor binding of Lrp, resulting in the Dam site becoming fully methylated in the daughter cells. Conversely, at a low, stochastic frequency within a population of cells, Lrp will bind to hemimethylated DNA following cell division, which will sequester this site from Dam. After a second round of DNA replication and cell division, this site can become fully unmethylated and Lrp will be stably bound to switch pilus expression into the OFF phase. This system, therefore, defines a heritable epigenetic switch for gene regulation.

To determine whether PVCA0063 undergoes a similar mechanism of regulation, we first determined whether Fur-mediated
undermethylation of PVCA0063 was responsive to the available iron concentration. Under iron-replete conditions when Fur should bind tightly, we find undermethylation of the Dam site in PVCA0063 (Fig. 5B). Furthermore, under iron-depleted conditions, when the affinity of Fur for its binding site is reduced, we find that the locus is fully methylated (Fig. 5B). The expression levels of VCA0063 (ptrB) were also assessed under these conditions by qRT-PCR, and the transcript levels were found to correlate directly with the methylation state of the promoter (Fig. 5C).

To determine whether Fur binding is inhibited by methylation

FIG 3 Characterization of intergenic undermethylated sites by MeSR digestion and qPCR. (A and B) Assays were assessed for reliable detection of undermethylation using gDNA from methylase-expressing and methylase-null mutants. (A) For MboI (solid line) and DpnI (dotted line), gDNA from Dam− and Dam+ E. coli strains were mixed in 10% increments from 100% to 0% Dam+ and the experimental values obtained (Measured) were compared to actual mixed ratios (Expected). (B) For MspI, gDNA from wild-type E7946 and a vchM mutant strain were mixed and analyzed as described for panel A. Data are shown as percent methylated relative to an uncut control reaction mixture. The slope of the linear regression ± standard error is indicated. Data are representative results from one of at least two independent experiments. (C to J) Characterization of Dam (C to G) and VchM (H to J) undermethylated intergenic sites using MeSR digestion and qPCR of gDNA from cultures grown to mid-exponential (optical density at 600 nm [OD600] of ~0.5) or stationary (OD600, ~3.0) phase. Data are shown as percent methylated relative to an uncut control reaction mixture, and significance was determined by one-sampled Student’s t test to determine if means were significantly different from 1. Each data point in panels C to J represents an independent biological replicate, and a horizontal line represents the median of each sample. *, P < 0.05; **, P < 0.01; ***, P < 0.001.
of \( P_{VCA0063} \) (as observed for the \( sciI \) and \( pap \) promoters in \( E. coli \)), we performed EMSAs. Fur can generate two independent shifts (higher- and lower-mobility complexes) in EMSAs due to the structure of the Fur box as two overlapping heptamer repeats (23). Using an unmethylated and an \( in \) \( vitro \) Dam-methylated probe of the VCA0063 promoter, we show that formation of the lower-mobility complex is inhibited when \( P_{VCA0063} \) is fully methylated (Fig. 5D). Full methylation of probes was confirmed by restriction digest of probes with the MeSR MboI and the MeDR DpnI (data not shown). It was previously shown that \( P_{PSodA} \) was bound by Fur, so a probe of this promoter was included in these assays as a positive control, while a probe of the \( aphA \) promoter was included as a negative control since this promoter is not predicted to interact with Fur (Fig. 5D).

Since we see that the mobility of Fur in EMSAs is altered by methylation of the VCA0063 promoter, it is possible that Fur binding and methylation by Dam constitute a heritable epigenetic switch that regulates this locus. There are no DNA demethylases in \( V. cholerae \), so transition from the methylated state to the unmethylated state competition for Dam and Fur binding at this site and is not the result of an epigenetic switch.

of \( P_{VCA0063} \) (as observed for the \( sciI \) and \( pap \) promoters in \( E. coli \)), we performed EMSAs. Fur can generate two independent shifts (higher- and lower-mobility complexes) in EMSAs due to the structure of the Fur box as two overlapping heptamer repeats (23). Using an unmethylated and an \( in \) \( vitro \) Dam-methylated probe of the VCA0063 promoter, we show that formation of the lower-mobility complex is inhibited when \( P_{VCA0063} \) is fully methylated (Fig. 5D). Full methylation of probes was confirmed by restriction digest of probes with the MeSR MboI and the MeDR DpnI (data not shown). It was previously shown that \( P_{PSodA} \) was bound by Fur, so a probe of this promoter was included in these assays as a positive control, while a probe of the \( aphA \) promoter was included as a negative control since this promoter is not predicted to interact with Fur (Fig. 5D).

Since we see that the mobility of Fur in EMSAs is altered by methylation of the VCA0063 promoter, it is possible that Fur binding and methylation by Dam constitute a heritable epigenetic switch that regulates this locus. There are no DNA demethylases in \( V. cholerae \), so transition from the methylated state to the unmethylated state competition for Dam and Fur binding at this site and is not the result of an epigenetic switch.
Overlapping RpiR binding and Dam methylation do not mediate an epigenetic switch. Next, we sought to determine whether VC1784, which encodes the sialidase NanH in *V. cholerae*, is controlled by an epigenetic switch. It was previously shown that the growth of *V. cholerae* in medium containing sialic acid (Neu5Ac) as the sole carbon source induces the expression of this locus (19). To confirm this result and determine how the methylation pattern at PVC1784 changes when this locus is induced, we grew *V. cholerae* in M9 medium with either glucose or Neu5Ac as the sole carbon source. When grown with sialic acid, we find that VC1784 is induced transcriptionally and that the promoter becomes fully methylated (Fig. 6A and B). Since we have already determined that RpiR is responsible for the undermethylation of this promoter, this indicates that this transcription factor is a repressor of VC1784 and that Neu5Ac may limit the ability of RpiR to bind to PVC1784 (Fig. 4A).

Since we have defined a condition when the locus is fully expressed and methylated, as well as a condition when the locus is repressed and undermethylated, we can determine whether a heritable epigenetic switch controls the expression of this locus, as was done for the VCA0063 promoter. To that end, we took cells grown under inducing conditions (M9 plus Neu5Ac), which fully methylates the promoter, and subsequently grew them for multiple generations in medium with glucose as the sole carbon source, to determine whether the methylation state of this promoter is heritable. When we performed this analysis, we found that the promoter quickly became undermethylated following growth under repressive conditions (~4
undermethylated Dam sites between PL and PR, which are likely
This is especially interesting for the K139 prophage, which has two
ulate (5). Furthermore, the location of the undermethylated site
provides the first direct evidence for a DNA-protein interaction
vents the MTase from accessing its target site, and thus, our study
been confirmed. Binding of transcription factors at these sites pre-
ble for the observed methylation pattern for five of these sites. All
regions, and we have identified the transcription factors responsi-

FIG 6 P_{VC1784} is not controlled by an epigenetic switch. (A and B) Cells were grown in M9 medium with glucose or sialic acid (Neu5Ac) as the sole carbon source, and the methylation level of P_{VC1784} was determined by MeSR digestion and qPCR (A), while the expression level of VC1784 was determined by qRT-PCR (B). Significance was determined by two-tailed Student’s t test (A) and Mann-Whitney (B). (C) The heritability of P_{VC1784} methylation was tested by taking cells grown in M9 plus Neu5Ac (fully methylated) and incubating them for the indicated number of generations in M9 medium with glucose as the sole carbon source. (D) Heterogeneity in the methylation of the VC1784 promoter was assessed among single-colony isolates by MeSR digestion and qPCR. Each data point represents an independent biological replicate, and a horizontal line represents the median of each sample. *, P < 0.05; ***, P < 0.001.

generations) (Fig. 6C). Furthermore, we did not see heterogeneity in the methylation of this site among single-colony isolates when cells were grown under inducing conditions and subsequently plated and grown in repressive medium (Fig. 6D). Thus, this indicates that, similar to overlapping Fur and Dam sites, overlapping RpiR and Dam methylation do not constitute an epigenetic switch at this site (Fig. 6C and D).

DISCUSSION
In this study, we have described a modified high-throughput method for identifying unmethylated sites. To our knowledge, this is the first study to use second-generation sequencing technology to identify genome-wide undermethylation in a bacterium. Using this approach, we identified the undermethylated sites for Dam and VchM, the two known MTases in V. cholerae.

Many of these undermethylated sites occurred in intergenic regions, and we have identified the transcription factors responsible for the observed methylation pattern for five of these sites. All of these transcription factors were predicted to regulate the identified undermethylated promoters; however, none had previously been confirmed. Binding of transcription factors at these sites prevents the MTase from accessing its target site, and thus, our study provides the first direct evidence for a DNA-protein interaction between these DNA-binding proteins and the promoters they regulate (5). Furthermore, the location of the undermethylated site indicates the location where the transcription factor is bound. This is especially interesting for the K139 prophage, which has two undermethylated Dam sites between P_{L} and P_{R}, which are likely the result of cI repressor binding. Using the location of the undermethylated sites, we identified putative cI operator sites for the K139 prophage. It is tempting to speculate that Dam methylation of these sites affects lytic induction of the phage; however, this hypothesis remains to be tested.

Interestingly, four of the undermethylated intergenic sites are located in operons involved in uptake/catabolism of carbon sources. P_{VC1778} is in a sialic acid utilization operon, P_{VC0286/7} is in a gluconate utilization operon, P_{VC1280/1} is in a chitobiose utilization operon, and P_{VC1358/9} is upstream from 6-phospho-beta-glucosidase (19, 24). A fifth site (P_{VC1006/8}) is upstream from a heme utilization operon, indicating that five undermethylated sites are in regions of the genome involved in scavenging of essential nutrients (iron and carbon) (20). For two of these sites (P_{VC0063} and P_{VC1784}), we show that undermethylation correlates with reduced transcription of the locus, suggesting that under normal growth in rich medium, these operons are repressed. It is likely that the other three transcription factors that promote undermethylation are also repressors, because two of these, VC1286 and VC1557, are homologs of the LacI repressor, while VC0289 is a homolog of the gluconate transcriptional repressor from E. coli. Thus, these undermethylated sites have revealed that the iron- and carbon-scavenging loci they reside in are likely tightly repressed during growth in LB. This is consistent with a report indicating that the primary carbon source during growth in LB is amino acids (25). Additionally, it is known that LB contains ~17 μM iron, making it an iron-rich growth medium (26).

Dam methylation has previously been shown to control gene expression via an epigenetic switch in other bacteria (2). Thus, we determined whether two of the undermethylated Dam sites we identified by methyl HTM-seq promoted a Dam-mediated epigenetic switch. We found that in P_{VC0063}, methylation reduced the affinity for Fur binding at this site. However, we found no evidence that the methylation state of cells was heritable, indicating that methylation likely did not constitute an epigenetic switch at this site. A similar analysis of P_{VC1784} also showed that this promoter was not subject to regulation by an epigenetic switch.

It was recently suggested that overlapping Dam and Fur mediated an epigenetic switch that controls the expression of a type VI secretion system in a pathogenic strain of E. coli (21). In that study, Fur binding was shown to prevent methylation of a Dam site in the sci promoter both in vitro and in vivo. Additionally, methylation of this Dam site was shown to reduce the affinity of Fur for this promoter in EMSAs. This was used to suggest that this regulatory arrangement mediates an epigenetic switch for the expression of the type VI secretion apparatus. What was not analyzed in that study, however, was a heritability of the methylation state following cell division to prove that regulation of this site truly occurred via an epigenetic switch (21). In our study, we found a very similar regulatory arrangement at P_{VC1006/8}. While we found a difference in the mobility of Fur in EMSAs for the methylated and unmethylated VCA0063 promoter, we did not find evidence for the methylation state of cells being heritable, indicating that this arrange-
ment likely does not constitute an epigenetic switch. Thus, it may be pertinent to further test the role of overlapping Fur and Dam in mediating an epigenetic switch for the type VI secretion system in pathogenic E. coli.

In E. coli, the expression of Pap pili is controlled by an epigenetic switch, and phase OFF cells are selected for during growth in rich medium in vitro, likely due to the cost of production of this pilus, while phase ON cells are selected for during infection, likely due to the importance of this pilus for virulence (27–29). The E7946 V. cholerae strain used in this study is a clinical isolate that has been subsequently passaged an estimated five times by colony purification on LB agar and expansion in LB broth, which represents ~170 generations in vitro. During this extensive growth in vitro, we would expect any analogous phase ON virulence loci that impose a fitness cost, and whose expression requires methylation at their promoters, to shift to the OFF phase (undermethylation). Since we did not identify any such loci, we conclude that there is likely not a virulence locus regulated by a Dam-dependent epigenetic switch in V. cholerae.

This genome-wide screen revealed that there were relatively few intergenic undermethylated Dam sites located in the V. cholerae genome. This is in contrast to E. coli, which has been shown to have 23 undermethylated intergenic Dam sites (5). Additionally, in pathogenic E. coli, some undermethylated Dam sites have been shown to promote an epigenetic switch in gene regulation; however, in V. cholerae, we found no evidence for a similar mechanism of regulation at the undermethylated sites identified (22, 30). Expression of the Dam methylase is essential in V. cholerae, while in E. coli, Dam mutant strains are still viable (6,31). Thus, it is tempting to speculate that the essential nature of Dam methylase activity in V. cholerae has evolutionarily limited its use in gene regulation.

Dam methylation in promoters has also been shown to affect transcription in multiple systems independent of mediating an epigenetic switch. These include control of transposase expression and transcription in multiple systems independent of mediating an epigenetic switch. These include control of transposase expression and transcription in multiple systems independent of mediating an epigenetic switch.

In conclusion, we have identified the undermethylated sites in the genome of a clinically relevant V. cholerae strain that may play a role in regulating gene expression in V. cholerae, which will be the focus of future studies.

ACKNOWLEDGMENTS

This work was supported by U.S. National Institutes of Health grant AI055058 (A.C.), and A.C. is a Howard Hughes Medical Institute investigator.

REFERENCES


tive/dependent restriction enzymes and real-time PCR. Epigenetics 1:146–152.


