



# Integration of the *Salmonella* Typhimurium Methylome and Transcriptome Reveals That DNA Methylation and Transcriptional Regulation Are Largely Decoupled under Virulence-Related Conditions

Jeffrey S. Bourgeois,<sup>a,b</sup> Caroline E. Anderson,<sup>a</sup> Liuyang Wang,<sup>a</sup> Jennifer L. Modliszewski,<sup>c</sup> Wei Chen,<sup>c</sup> Benjamin H. Schott,<sup>a,b</sup> Nicolas Devos,<sup>c</sup> Dennis C. Ko<sup>a,b,d</sup>

<sup>a</sup>Department of Molecular Genetics and Microbiology, School of Medicine, Duke University, Durham, North Carolina, USA

<sup>b</sup>University Program in Genetics and Genomics, Duke University, Durham, North Carolina, USA

<sup>c</sup>Center for Genomics and Computational Biology, Duke University, Durham, North Carolina, USA

<sup>d</sup>Division of Infectious Diseases, Department of Medicine, School of Medicine, Duke University, Durham, North Carolina, USA

**ABSTRACT** Despite being in a golden age of bacterial epigenomics, little work has systematically examined the plasticity and functional impacts of the bacterial DNA methylome. Here, we leveraged single-molecule, real-time sequencing (SMRT-seq) to examine the m<sup>6</sup>A DNA methylome of two *Salmonella enterica* serovar Typhimurium strains: 14028s and a  $\Delta metJ$  mutant with derepressed methionine metabolism, grown in Luria broth or medium that simulates the intracellular environment. We found that the methylome is remarkably static: >95% of adenosine bases retain their methylation status across conditions. Integration of methylation with transcriptomic data revealed limited correlation between changes in methylation and gene expression. Further, examination of the transcriptome in  $\Delta yhdJ$  bacteria lacking the m<sup>6</sup>A methylase with the most dynamic methylation pattern in our data set revealed little evidence of YhdJ-mediated gene regulation. Curiously, despite G(m<sup>6</sup>A)TC motifs being particularly resistant to change across conditions, incorporating *dam* mutants into our analyses revealed two examples where changes in methylation and transcription may be linked across conditions. This includes the novel finding that the  $\Delta metJ$  motility defect may be partially driven by hypermethylation of the chemotaxis gene *tsr*. Together, these data redefine the *S. Typhimurium* epigenome as a highly stable system that has rare but important roles in transcriptional regulation. Incorporating these lessons into future studies will be critical as we progress through the epigenomic era.

**IMPORTANCE** While recent breakthroughs have enabled intense study of bacterial DNA modifications, limitations in current work have potentiated a surprisingly untested narrative that DNA methylation is a common mechanism of the bacterial response to environmental conditions. Essentially, whether epigenetic regulation of bacterial transcription is a common, generalizable phenomenon is a critical unanswered question that we address here. We found that most DNA methylation is static in *Salmonella enterica* serovar Typhimurium, even when the bacteria are grown under dramatically different conditions that cause broad changes in the transcriptome. Further, even when the methylation of individual bases change, these changes generally do not correlate with changes in gene expression. Finally, we demonstrate methods by which data can be stratified in order to identify coupled changes in methylation and gene expression.

**KEYWORDS** DNA methylation, *Salmonella*, gene regulation, m<sup>6</sup>A, methylome, transcription

**Invited Editor** Jay Hinton, University of Liverpool

**Editor** Joerg Vogel, University of Würzburg

**Copyright** © 2022 Bourgeois et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Dennis C. Ko, [dennis.ko@duke.edu](mailto:dennis.ko@duke.edu).

The authors declare no conflict of interest.

**Received** 16 November 2021

**Accepted** 11 May 2022

**Published** 6 June 2022

Until recently, systematically understanding how the bacterial DNA methylome affects physiology has been an unachievable task. Unlike eukaryotes where m<sup>5</sup>C DNA methylation is highly abundant and can be detected with bisulfate sequencing (1), bacterial genomes primarily house m<sup>6</sup>A methylation, which has historically been difficult to detect. Despite this technological hurdle, many studies over the last several decades have successfully uncovered roles for DNA methylation in contexts of restriction-modification systems (reviewed in reference 2), as well as for “orphan” methylases (particularly the Dam methylase) in DNA repair (3–9), DNA/bacterial replication and viability (10–19), *agn43* phase variation (20), LPS modifications (21–25), phage defense (21, 26, 27), mating (28, 29), fimbria formation (30, 31), antibiotic resistance (32), hypoxia survival (33), motility (17, 23, 31, 34), and other virulence-related processes (8–10, 16–19, 23, 31, 34–39). While orphan methylases were originally thought to regulate bacterial physiology, whereas restriction-modification systems targeted foreign DNA, recent work on “phasevarions” has shown that restriction-modification systems can indeed have a dramatic impact on the genome (reviewed in reference 40). A more complete history of associations between methylases and phenotypes can be found in recent reviews (41–43).

While early epigenome studies are valuable for the insights they provide, they depended on low-throughput and relatively blunt approaches (e.g., restriction enzyme digests paired with Southern blotting to infer methylation). These approaches could not be leveraged to address whether and how genome-wide changes in DNA methylation associate with changes to cellular processes. However, the discovery that sequencing data from Pac-Bio single-molecule, real-time sequencing (SMRT-seq) (44) and Oxford Nanopore sequencing (45, 46) systems can be repurposed to detect m<sup>6</sup>A has heralded a golden age of bacterial DNA methylomics. These technological breakthroughs were rapidly applied to cataloging bacterial methylomes, many of which have been deposited in publicly available databases such as REBASE (47). However, we have only recently seen the power of these third-generation sequencing technologies applied to connect DNA methylation to cellular phenotypes. For instance, a recent paper utilized SMRT-seq to identify specific changes in G(m<sup>6</sup>A)TC patterns within the *opvAB* promoter that are highly present in the population following phage insult (48), building on previous phenotypic observations (21). Other groups have leveraged comparative epigenomics to examine methylation patterns across isolates and identify potentially important trends in methylation (35, 49, 50). Thus, there is immense potential for SMRT-seq to identify how methylation correlates with impactful biology.

Despite these advances, we note that few studies have leveraged SMRT-seq to understand how methylation itself changes under different environmental pressures. Instead, the studies listed above typically examine methylomes under a single condition (typically late stationary-phase growth) to infer where methylation *can* happen, with notable exception (11, 32, 51–53). While informative, these approaches may not represent the methylation status of bacteria at growth phases typically studied in bacteriology and thus may have limited ability to integrate into the broader microbiological literature and with transcriptomic data sets. A related shortcoming of many methylation studies is that methylation sites in promoters are often reported as evidence of methylation-mediated regulation, without determining whether methylation is dynamic under relevant environmental changes or testing whether disrupting methylation at those sites impacts transcription. Curiously, while a number of classical approaches have identified examples of methylation at specific sites regulating gene expression (e.g., *pap* [30, 54, 55], *opvAB* [21], *agn43* [56, 57], *gtr* [22], the *std* operons [58], *dnaA* [12], *traJ* [29], and *sci1* [36]), we are unaware of any methylation site originally identified using genome-wide approaches that has been confirmed to impact gene expression. This disconnect between the technological advances in methylomics and the relatively modest conceptual advances in the field make it clear that the use of third-generation sequencing technologies to interrogate the DNA methylome is still in

its infancy and that the important question of the generalizability of epigenetic regulation of transcription in bacteria should be addressed.

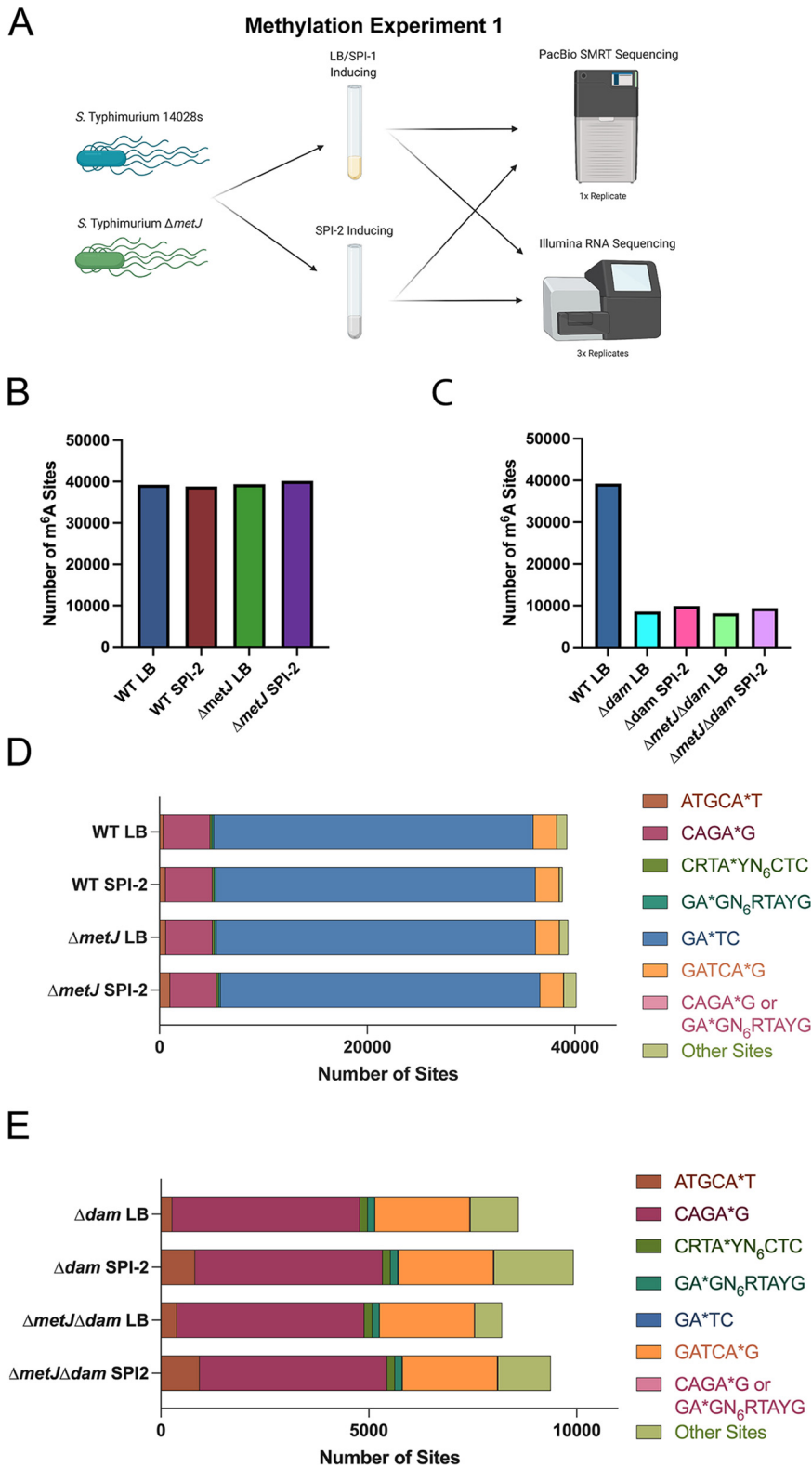
Here, we perform a series of SMRT-seq and transcriptome sequencing (RNA-seq) experiments to understand the role of DNA methylation in regulating *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*) gene expression under environmental conditions critical for *Salmonella* virulence. Specifically, we studied conditions that activate the motility and *Salmonella* Pathogenicity Island-1 (SPI-1) pathways (growth in Luria broth [LB] until late log phase) and conditions that activate the *Salmonella* Pathogenicity Island-2 (SPI-2) pathways (growth in low phosphate and magnesium [LPM] medium [59] until late log phase). Since methionine metabolism is intimately connected to methylation, we also examined the changes in methylation associated with derepressed methionine metabolism using a  $\Delta metJ$  mutant. In general, we found that DNA methylation is mostly stable across conditions and is broadly decoupled from gene expression changes. This study redefines our understanding of the *S. Typhimurium* epigenome, provides multiple epigenomic data sets that can be incorporated into future work to identify rare instances where methylation changes are coupled with transcription, and presents a basic blueprint for carrying out future methylomic studies with high reproducibility.

## RESULTS

**A genome-wide screen to understand how growth conditions and methionine metabolism impact m<sup>6</sup>A DNA methylation.** While previous work on bacterial DNA methylation has largely focused either on global DNA methylation patterns under a single condition or on how methylation of a single motif changes under different conditions, we sought to examine how the entire *S. Typhimurium* m<sup>6</sup>A DNA methylome changes under four biologically relevant conditions (Fig. 1A). We examined aerobic growth in LB to late exponential phase (optical density at 600 nm [OD<sub>600</sub>] ~ 1.5 to 2.0), which induces expression of flagellar genes and the genes in the *Salmonella* Pathogenicity Island-1 (SPI-1)—including the type III secretion system used during host cell invasion (60). The second condition cultured bacteria in a minimal media used to induce expression of genes in the *Salmonella* Pathogenicity Island-2 (SPI-2) (59)—which include the type III secretion system turned on in the host cell to promote *Salmonella* vacuolar survival (61, 62). The third and fourth conditions repeated growth in these media but used a methionine metabolism mutant *S. Typhimurium* strain ( $\Delta metJ$ ). The MetJ protein represses expression of methionine metabolism genes (63). Thus,  $\Delta metJ$  bacteria have deregulated methionine metabolism and accumulation of methionine and related metabolites, including metabolites directly related to methylation processes such as the universal methyl-donor *S*-adenosylmethionine (SAM) (64), and the methyltransferase-inhibiting metabolites methylthioadenosine (65–67) and *S*-adenosylhomocysteine (68–72). Of note, we have not only previously confirmed increased abundances of both SAM and methylthioadenosine in the  $\Delta metJ$  mutant but also demonstrated that the  $\Delta metJ$  mutant has attenuated SPI-1 secretion, motility, and virulence (73), and we had previously hypothesized that these effects could be mediated through aberrant methylation.

To analyze the DNA methylome, we performed a PacBio SMRT-seq experiment (here called “methylation experiment 1”) in a biological singlet (as has been common in the field and as we comment on below) to identify whether any changes in methylation could be observed. In this experiment, we also included  $\Delta dam$  and  $\Delta dam \Delta metJ$  mutants, which lack G(m<sup>6</sup>A)TC (henceforth an asterisk [\*] will denote the adenosine that is m<sup>6</sup>A modified, e.g., “GA\*TC”) methylation grown under SPI-1 and SPI-2 conditions. This allowed us to confirm that our pipeline could adequately detect changes in methylation. These eight conditions were split across two PacBio SMRT cells. Thus, these conditions enabled comparison of the *S. Typhimurium* DNA methylome under the two conditions most critical for *Salmonella* virulence (SPI-1 and SPI-2 induction), following perturbation of methionine metabolism, and with a control condition ablating the primary DNA methyltransferase.

In total, this experiment defined the methylation status of 61,704 adenosine bases (GEO [GSE185578](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185578)); however, methylation status of some bases under certain conditions



**FIG 1** Genome-wide analysis of m<sup>6</sup>A DNA methylation under various conditions. (A) Schematic of methylation experiment 1. Wild-type *S. Typhimurium* (strain 14028s) and an isogenic  $\Delta metJ$  strain were cultured in LB or SPI-2-inducing media, and DNA was collected for SMRT-seq. Bacteria grown under identical conditions were harvested for RNA-sequencing. (B and C) The total number of m<sup>6</sup>A bases observed across conditions does not dramatically change in wild-type and  $\Delta metJ$  bacteria (B) but does change dramatically in  $\Delta dam$  bacteria (C). (D) Analysis of motifs methylated reveals only the

(Continued on next page)

could not be determined since coverage was below 50 $\times$ . Thus, we restricted our analysis to 51,177 bases in which the methylation status could be adequately determined for all conditions tested. These bases span both the *S. Typhimurium* genome and virulence plasmid.

To compare methylation across conditions, we called methylation in two ways. First, we assigned each base a “percent methylated” value, which considered the percentage of reads for each base that were counted as methylated compared to the total number of reads (see Data Set S1). We also examined the data as a binary value in which we considered bases either methylated (if any methylation was detected) or unmethylated (see Data Set S2). Using this binary analysis, we observed that there were similar, but subtly different, amounts of m<sup>6</sup>A methylation across wild-type (WT) and  $\Delta metJ$  bacteria grown under LB and SPI-2 conditions (WT LB, 39,240 bases; WT SPI-2, 38,827 bases;  $\Delta metJ$  SPI-1, 39,352 bases;  $\Delta metJ$  SPI-2, 40,145 bases) (Fig. 1B), but that, as expected, the  $\Delta dam$  mutation reduced total methylation substantially (Fig. 1C).

**ATGCAT motifs are frequently differentially methylated across conditions.** We next examined how these bases were distributed across different methylation motifs. This analysis detected methylation at motifs that had been previously detected in *S. Typhimurium* (74), though we were able to detect an additional motif, CRTA\*Y<sub>n</sub>CTC, which appears to be the reverse complement of GA\*GN<sub>6</sub>RTAYG. Notably, two motifs (CAGA\*G and GA\*GN<sub>6</sub>RTAYG) cannot always be distinguished, and so we included bases (~20 bases per condition) that matched both motifs in a separate category: “CAGA\*G or GA\*GN<sub>6</sub>RYAYG.” Bases that did not map to any known motif were listed as “Other.” As with the total amount of m<sup>6</sup>A methylation, we found that our four main conditions had subtle differences in the total numbers of most motifs (Fig. 1D).

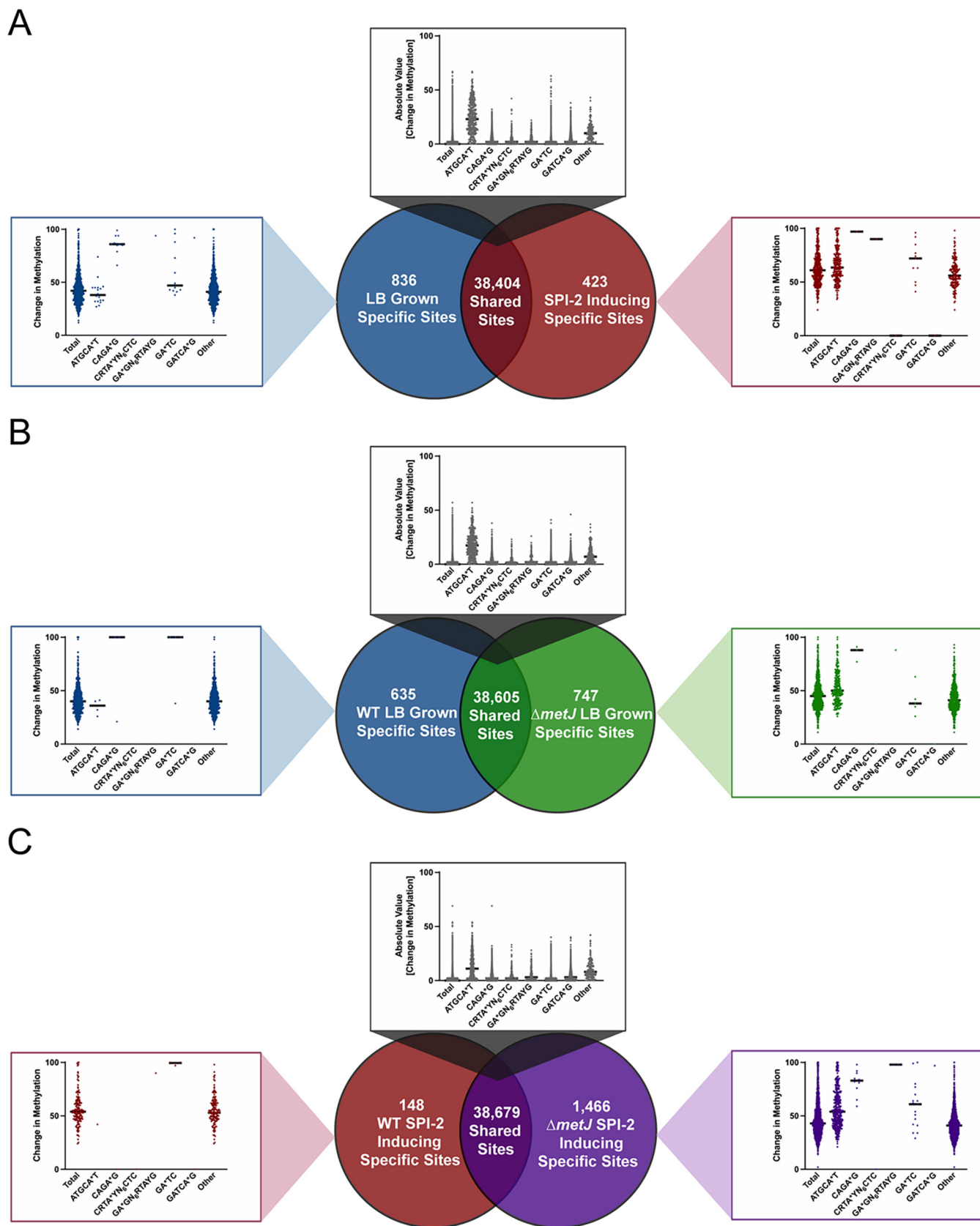
The most notable change in motif abundance occurred at the ATGCA\*T motif, which is methylated by the YhdJ methylase (75). We observed more ATGCA\*T methylation in bacteria grown under SPI-2-inducing conditions ( $P < 0.00001$ , chi-square test) or in  $\Delta metJ$  bacteria ( $P < 0.00001$ , chi-square test), with the highest ATGCA\*T methylation present in  $\Delta metJ$  bacteria grown under SPI-2-inducing conditions. We also observed variation in the number of bases that mapped to the “Other” category. In contrast, we observed very little change in the total amount of GA\*TC methylation (methylated by Dam) across these physiologically relevant conditions, though the deletion of *dam* resulted in almost complete ablation of methylation at the GATC motif (Fig. 1E).

To examine the methylation changes under LB versus SPI-2 inducing conditions with wild-type and  $\Delta metJ$  *S. Typhimurium*, we compared binary methylation at each individual base to identify differentially methylated bases (bases that were called methylated in one condition but not another). While each condition had a few hundred to over a thousand bases that were not methylated in their opposing group, the vast majority of bases in this study (>38,000) were shared across these comparisons (Fig. 2A to C; Venn diagrams). This demonstrates that while the methylome is slightly responsive to the environment and methionine metabolism, it remains largely static across strikingly different conditions.

Having identified these differentially methylated bases by our binary analyses, we integrated our quantitative data. This is an important measurement as previous work has speculated that methylation impacts bistable gene expression (48, 76), and thus changes in the percentage of the population in which a given base is methylated could have implications on the percentage of the population expressing a given gene. For each differentially methylated base, we asked what the total change in methylation

#### FIG 1 Legend (Continued)

total number of ATGCA\*T and “other” sites (sites that do not map to one of the six motifs) changes dramatically across conditions. The roughly 20 sites that could not be distinguished between CAGA\*G or GA\*GN<sub>6</sub>RTAYG methylation are listed as “CAGA\*G or GA\*GN<sub>6</sub>RYAYG.” (E)  $\Delta dam$  mutation results in ablation of GATC methylation. For panels B through E, bases were only included in the analysis if the base could confidently be called methylated or unmethylated across the eight conditions.



**FIG 2** Integration of binary and quantitative analyses to understand differential methylation in *S. Typhimurium*. (A to C) Quantification of shared and unique methylated sites between wild-type *S. Typhimurium* grown in LB and SPI-2-inducing media (A), WT and  $\Delta metJ$  bacteria grown in LB (B), and WT (Continued on next page)



was across the two conditions (Fig. 2; see graphs in panels A to C). The total median shift in the percent methylation varied by condition, but fell between 43 and 53%, suggesting that most bases go from unmethylated in one condition, to about half the copies of the genome having methylation at that site in the other. Notably, most of the “shared” bases that demonstrated methylation under both conditions demonstrated no quantitative change (median = 0%). However, again the exception was ATGCA\*T, where the median shift among shared bases remained relatively high (11 to 23%, depending on the condition).

To test for enrichment of motifs among differentially methylated bases, we compared the frequency of each of the six motifs tested above in the differentially methylated sites against the frequency observed in the entire condition. This analysis revealed that among the uniquely SPI-2-induced methylated bases, we observed 37 times more differentially methylated ATGCA\*T sites than expected. Similarly, sites methylated in  $\Delta metJ$  *S. Typhimurium*, but not wild-type bacteria, grown under both LB and SPI-2-inducing conditions are also dramatically enriched for YhdJ-mediated methylation (20- and 11-fold enrichment, accordingly) (see Fig. S1A in the supplemental material). Surprisingly, all other motifs were either present at similar or dramatically lower abundance among differentially methylated sites than expected by chance, though we did note significant enrichment of “other” motifs among differentially methylated bases (20- to 100-fold enrichment, depending on the condition).

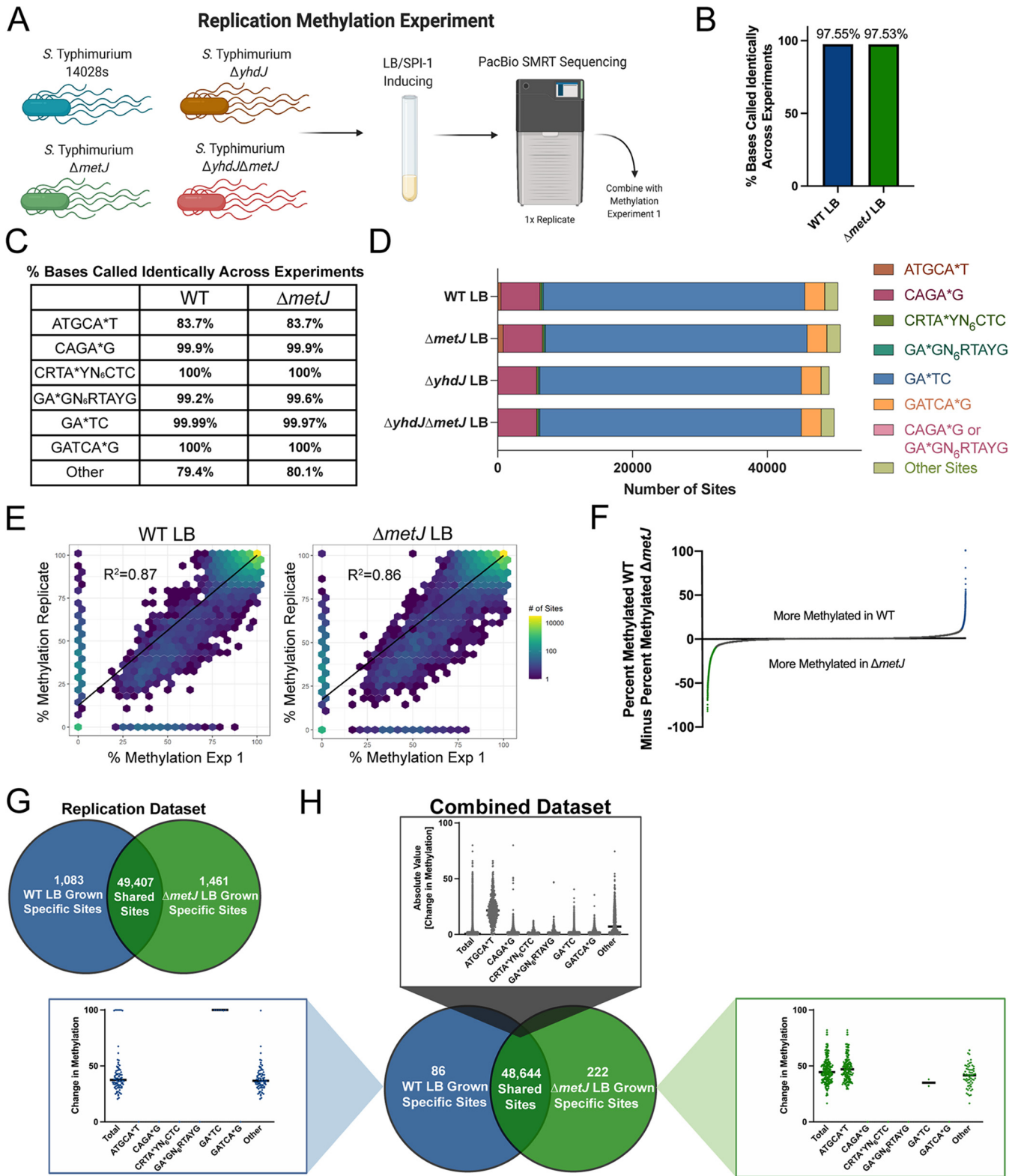
**A replication experiment demonstrates that SMRT-seq is highly reproducible and confirms differential methylation is predominantly driven by YhdJ.** To confirm our findings that DNA methylation was largely stable among our conditions with the exception of ATGCA\*T sites, we repeated our SMRT-seq experiment with wild-type and  $\Delta metJ$  bacteria grown in LB (Fig. 3A, replication methylation experiment). Further, to confirm that the significant enrichment in ATGCA\*T methylation in  $\Delta metJ$  bacteria that we observed above was due to YhdJ, we also sequenced  $\Delta yhdJ$  and  $\Delta yhdJ \Delta metJ$  *S. Typhimurium* grown in LB. Of note, although we sequenced eight samples across two SMRT cells in methylation experiment 1, we sequenced these four samples on two SMRT cells, significantly increasing our sequencing depth. The resulting data set called the methylation status of 60,502 bases under at least one condition; strikingly, 60,501 of these bases were confidently called in all four conditions (GEO [GSE185501](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185501)).

Analysis of the two methylomic data sets using our binary assessment revealed that ~97.5% of bases replicated their methylation status across experiments, demonstrating that our results were highly reproducible (Fig. 3B; see Data Set S2). Of note, sites assigned to ATGCA\*T or “other” motifs may include significantly more miscalled methyl bases, as only ~80% (79.3% of wild-type strains and 80.1% of  $\Delta metJ$  strains) “other” bases replicated their methylation status across experiments (Fig. 3C). Alternatively, as we see these bases are more dynamic across our conditions, these motifs may represent biologically meaningful sources of variation across experiments. Importantly, we again observed that  $\Delta metJ$  bacteria have increased ATGCA\*T methylation and found 0 methylated ATGCA\*T sites in  $\Delta yhdJ$  and  $\Delta yhdJ \Delta metJ$  mutants, confirming that YhdJ is the only ATGCA\*T methylase active in both bacterial strains (Fig. 3D).

Analysis of the two data sets using the quantitative measurement (see Data Set S3) revealed considerable replication across the data sets (wild type,  $R^2 = 0.87$ ;  $\Delta metJ$ ,  $R^2 = 0.86$ ) (Fig. 3E). Considering these experiments as separate biological replicates and using an arbitrary cutoff of 10% average differential methylation, we identified 2,528 sites (of 50,962 total sites; 4.96%) that were differentially methylated between wild-type and  $\Delta metJ$  bacteria using this quantitative method (Fig. 3F). A total of 881 of

## FIG 2 Legend (Continued)

and  $\Delta metJ$  bacteria grown in SPI-2-inducing media (C). Venn diagrams are based on binary measures of differential methylation. Sites identified by the binary analysis were examined in our quantitative data set in order to identify changes in the percent methylation. In the graphs, “Total” refers to all sites present in the relevant part of the Venn diagram, which were then broken down by motif. For motifs where no differentially methylated sites were present, a single dot is listed at 0%. For shared sites, the absolute value of the difference between bases are shown, and thus the numbers are agnostic to whether methylation is higher in either condition. Bars mark the median. For all panels, only bases that could be confidently called methylated or unmethylated under the eight conditions in Fig. 1 were considered.



**FIG 3** A replication screen reveals methylation is highly reproducible across SMRT-seq experiments but highlights the value of performing biological replicates. (A) Schematic for the replication methylation experiment. Wild-type *S. Typhimurium* (strain 14028s) or isogenic mutants were grown in LB media, and DNA was harvested for SMRT-seq. (B) Approximately 97% of all bases were called identically (methylated or unmethylated) in methylation experiment 1 and the replicate methylation experiments. (C) ATGCA\*T and "other" sites had lower rates of replication compared to other motifs. (D) Only ATGCA\*T and "other" sites (bases that do not map to one of the six motifs) change dramatically across tested conditions in the replication methylation experiment. No ATGCA\*T methylation was observed in  $\Delta yhdJ$  mutants. (E) The observed percent methylation at each base was reproducible across

(Continued on next page)



these sites were more methylated in wild-type bacteria, and 1,647 were more methylated in  $\Delta metJ$  bacteria.

Having assessed the reproducibility of SMRT-seq for both categorical and quantitative measures of methylation, we used our binary measurement to generate a combined data set containing bases which were (i) reliably detected in wild-type and  $\Delta metJ$  bacteria grown in LB in both experiments and (ii) were identically called methylated or unmethylated in both experiments. Using this data set (52,594 bases), we determined which differentially methylated bases repeated across the two studies. This number of bases is greater than the number of bases included in methylation experiment 1 analyses, since we no longer needed to exclude bases that did not reach sufficient coverage under either the SPI-2-inducing or  $\Delta dam$  conditions. While our data demonstrated that the vast majority of bases were called identically ( $\sim 97.5\%$ ; Fig. 3B), we found that a disproportional number of bases that were called differentially methylated in the pilot study failed to replicate in the replication study, and vice versa. This is in line with our observation that there was reduced reproducibility in our more dynamic ATGCA\*T and “other” motifs (Fig. 3C). In fact, while there were 1,382 bases called differentially methylated in the first experiment (Fig. 2B), and 2,544 bases called differentially methylated in the replicate study (Fig. 3G), only 308 differentially methylated bases were identified in the combined data set (Fig. 3H). Importantly, the overlap between these two replicates is much greater than expected by chance (3.7-fold enrichment;  $P < 0.0001$ , one-tailed binomial test), indicating that these biological replicates provide a high-confidence set of 308 differentially methylated sites, though some false positives likely remain. Our findings emphasize that while SMRT-seq calling of methylated bases is reliable, replication is especially important in examining bases that change between conditions.

Of note, the combined data set once again revealed enrichment of differentially methylated “other” sites. To understand these sites, we examined the 40 bases surrounding the 143 instances of “other” differential methylation in the combined data set using “Multiple Em for Motif Elicitation” (MEME) software (77). This identified a single significant motif (E value =  $6.1 \times 10^{-16}$ ), ACCWGG (see Fig. S1B). The same motif was identified among the 969 differentially methylated “other” sites between LB and SPI-2 grown bacteria (methylation experiment 1; E value =  $5.5 \times 10^{-206}$ ; see Fig. S1C). The ACCTGG motif has been reported on multiple *Salmonella* serovar entry pages on REBASE (47); however, curators note that this is almost certainly a miscall for the m<sup>5</sup>C motif CCWGG—methylated by Dcm. Across the combined data set, we found 33 instances of this motif (23% of all differentially methylated “other” sites). This leads us to hypothesize that this dynamic “other” category may be predominantly driven by changes in the flexible m<sup>5</sup>C methylome, which warrants further investigation using sequencing technologies better equipped to detect cytosine methylation.

**There is no association between the transcriptome and the genome-wide binary methylation analyses under the conditions tested here.** Canonically, changes in DNA methylation can lead to changes in transcription by enabling differential binding of transcription factors to genomic elements (reviewed in reference 41). However, studies that describe this in bacteria typically either (i) focus on single loci (see, for example, reference 48), or (ii) only speculate on direct mechanisms of transcriptional control based on methylase knockout experiments (see, for example, reference 35). No study

### FIG 3 Legend (Continued)

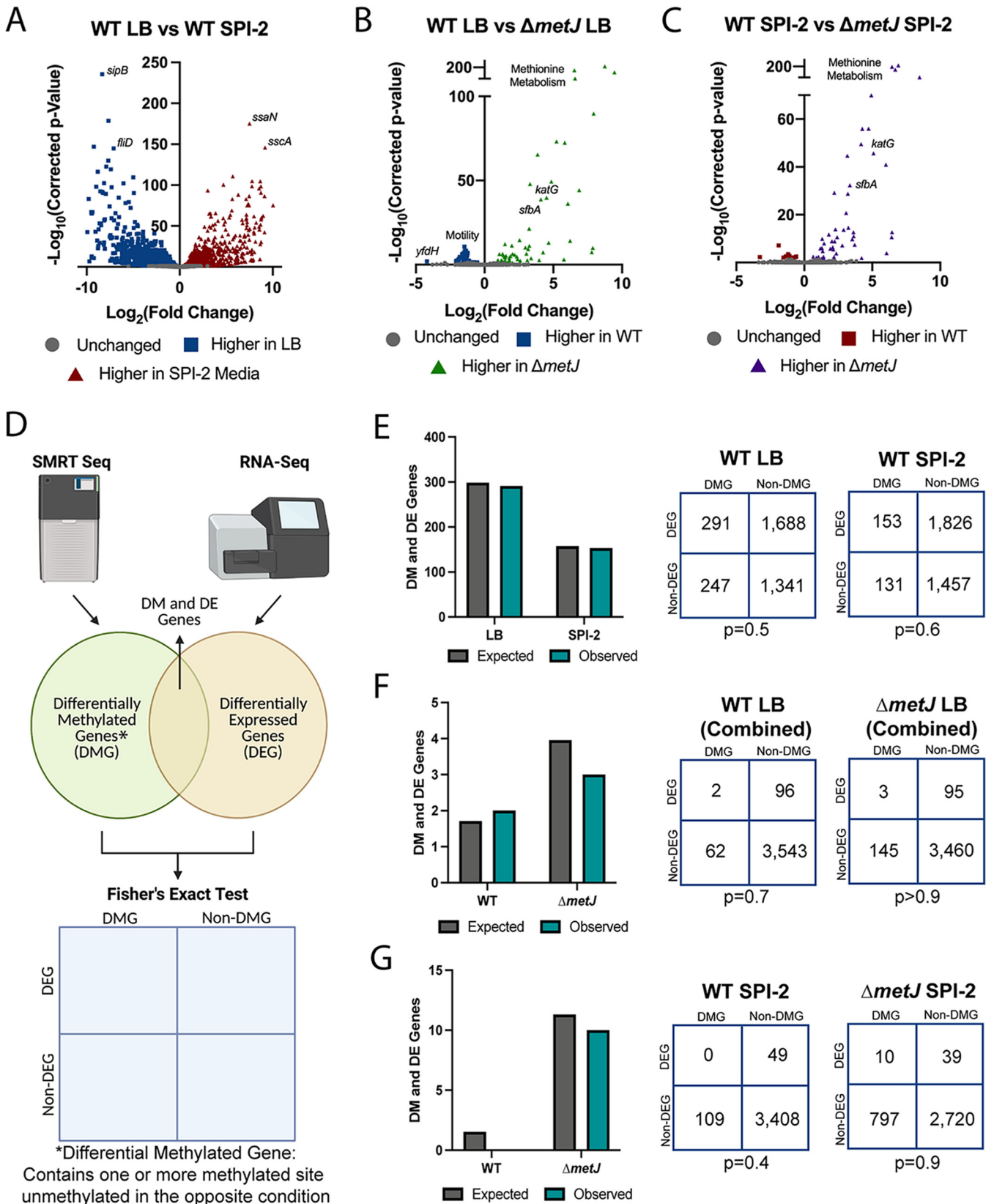
experiments. The color of the hexagon represents the number of bases that fall at that point on the axes.  $R^2$  values and trendlines represent the correlation across experiments. (F) Quantitative analysis reveals numerous sites are differentially methylated between wild-type and  $\Delta metJ$  bacteria. Each dot represents the mean percent methylation in wild-type bacteria across the two experiments subtracted by the mean methylation in  $\Delta metJ$  bacteria for each adenosine confidently called in both experiments. Blue and green dots mark bases where the mean difference is  $\geq 10\%$ . (G) Quantification of unique methylation sites in the replication experiment. For panels D to G, bases were only included in the analysis if the base could confidently be called methylated or unmethylated across conditions. (H) Venn diagram is based on binary measures of differential methylation in the combined data set. Sites identified by the binary analysis were examined in our quantitative data set in order to identify changes in the percent methylation. In the graphs, “Total” refers to all differentially methylated sites under that condition, and differentially methylated sites are then broken down by motif. For motifs where no differentially methylated sites were present, a single dot is listed at 0%. For shared sites, the absolute value of the difference between bases are shown and thus the numbers are agnostic to whether methylation is higher in either condition. Bars mark the median.

has directly examined whether differential methylation across the *S. Typhimurium* genome correlates with differential expression under biologically relevant conditions. We attempted to fill this gap in knowledge by performing RNA-seq on wild-type and  $\Delta metJ$  bacteria grown in LB and SPI-2-inducing media (GEO [GSE185072](#); see Data Set S4) and looking for correlations with our SMRT-seq data sets.

Prior to integrating our data sets, we confirmed our RNA-seq data matched previously observed trends. As expected, we identified many differentially expressed genes (DEGs; 2,639 DEGs at a false discovery rate [FDR] of  $<0.5$ ) between wild-type bacteria grown in LB versus SPI-2-inducing conditions (Fig. 4A). These DEGs included a variety of expected genes, including higher expression of SPI-1 genes (e.g., *sipB*, 322-fold induction) and flagellar genes (e.g., *fliD*, 138-fold induction) in LB and higher expression of SPI-2 genes (e.g., *ssaN* and *sscA*; 182- and 579-fold induction, respectively) in SPI-2-inducing media. Further, these data cluster with previous transcriptomic analyses of gene expression in LB and SPI-2-inducing conditions (78) by principal component analysis (PCA; 53.5% of variation between the two studies explained by media; see Fig. S1D). Analysis of  $\Delta metJ$  DEGs revealed a number of expected trends, specifically that during growth in LB and SPI-2,  $\Delta metJ$  bacteria show upregulation of methionine metabolism genes resulting from direct derepression of the metabolic pathway (Fig. 4B and C) (63). Of note, we also observed reduced motility gene expression which we had previously reported (73) but were surprised that, in contrast to our prior work, we observed a small reduction in *flhD* expression in the  $\Delta metJ$  mutant. However, we confirmed this result by qPCR and Western blotting (see Fig. S1E and F) and speculate that improved DNase treatment in this study likely explains this difference.

To integrate our differential expression data with our methylomics data, we considered genes that either (i) contained or (ii) were the closest gene to one or more binary differentially methylated sites (i.e., present at any level in one condition, absent in the other) to be “differentially methylated genes” (DMGs) (Fig. 4D). For each comparison, the status “differentially methylated” applied to the condition in which the methyl mark was present (e.g., when comparing LB-grown versus SPI-2-grown bacteria, an LB-grown DMG contains a methyl mark that is absent in SPI-2-grown bacteria). Using these criteria, we examined whether differentially methylated genes were more likely to be differentially expressed than predicted by chance. Strikingly, we did not observe enrichment of DEGs among our DMGs. The number of DEGs that were also DMGs under all comparisons was remarkably similar to the overlap of these categories expected by chance (Fisher  $P$  value  $> 0.05$  in all cases), suggesting the two phenomena are typically not associated at the genome-wide level across any of our comparisons with binary calling of DMGs (Fig. 4E to G). Most importantly, there was no evidence of enrichment comparing WT *S. Typhimurium* grown in LB (SPI-1 inducing) and SPI-2 conditions (Fig. 4E), indicating that cooccurrence of differential methylation and differential gene expression is not observed more frequently than is expected by chance in switching between these two critical growth conditions in *Salmonella* pathogenesis. Notably, we also did not observe correlations if we adjusted the statistical thresholds for differential expression (see Fig. S2A to D), or if we stratified our data by the differential expression direction of effect (see Fig. S2E and F), the genic location of the differential methylation (see Fig. S2G), or to specific motifs (see Fig. S2H and I).

**There is limited association between the transcriptome and the genome-wide quantitative methylation analyses under the conditions tested here.** In addition to these binary definitions to differential methylation, we examined whether there was enrichment of DEGs among DMGs defined by a difference in  $\geq 10\%$  methylation across conditions (Fig. 5A). Most of our binary observations replicated in this analysis (Fig. 5B and C), but we noted an association between DEGs and DMGs in wild-type bacteria grown in SPI-2-inducing media compared to  $\Delta metJ$  bacteria grown under the same conditions (Fig. 5D and Table 1). It is unclear why this condition is the exception to the general lack of association we observe, but it may suggest that methylases (particularly the *dam* methylase) and deregulated transcriptional machinery uniquely compete for access to these sites exclusively in minimal media. We also examined whether adjusting our thresholds for differential expression or limiting our search for differential to



**FIG 4** Differentially methylated genes by binary analysis are not enriched for transcriptomic changes (A to C) RNA-seq reveals transcriptomic changes between LB grown and SPI-2 media grown wild-type bacteria (A), wild-type and  $\Delta metJ$  bacteria grown in LB (B), and wild-type and  $\Delta metJ$  bacteria grown in SPI-2 media (C). Corrected *P* values generated by calculating the FDR. (D) Schematic of RNA-seq and SMRT-seq integration. Each gene was determined (Continued on next page)

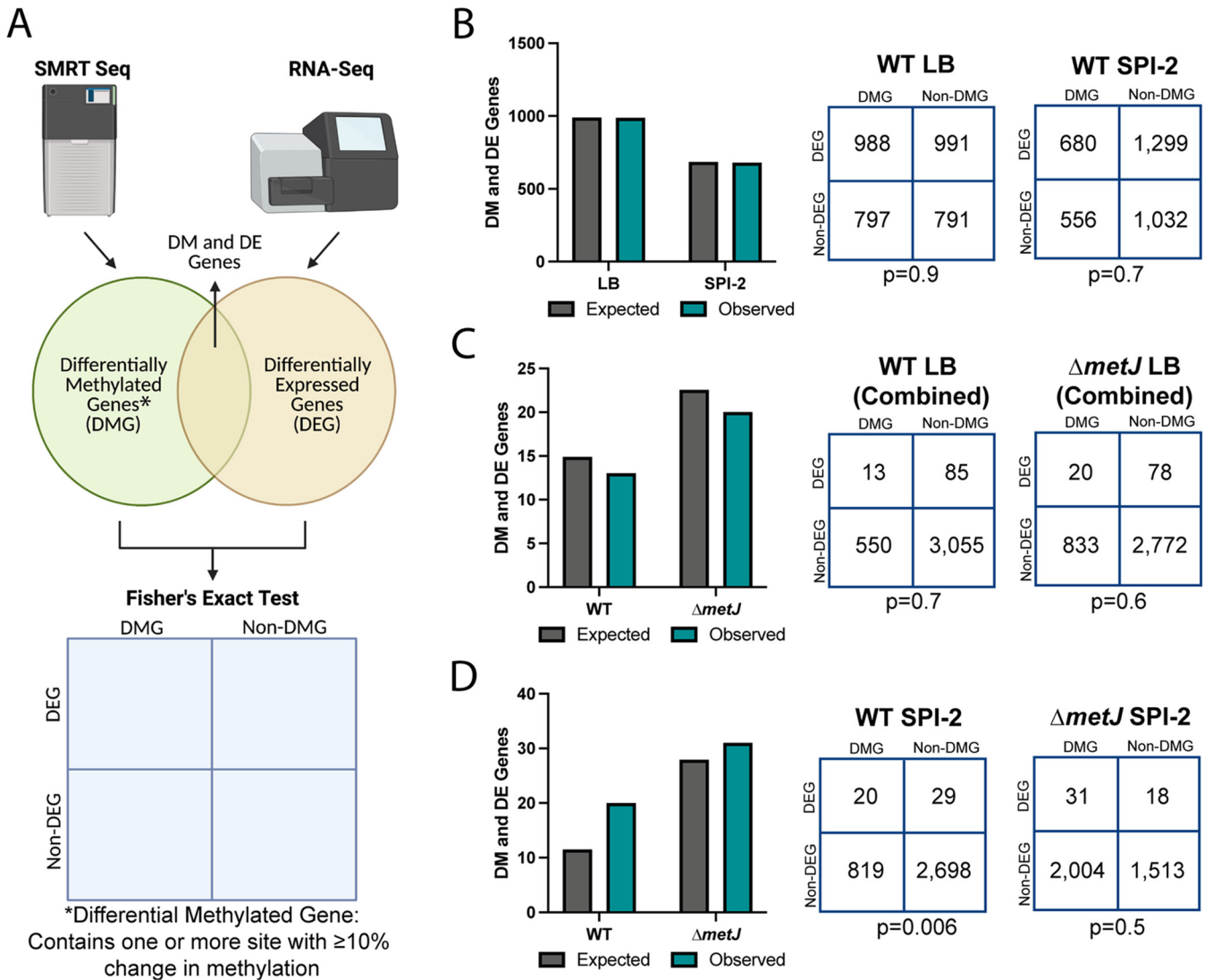
bases upstream of differentially expressed genes could reveal further correlations between expression and methylation, but no additional associations were found (see Fig. S2J to O). The association with sites in wild-type bacteria in SPI-2 media compared to  $\Delta metJ$  bacteria was still present with the more stringent DEG cutoff (see Fig. S2L) but was no longer present when analysis was restricted to upstream bases (see Fig. S2O). Finally, we leveraged the quantitative data set to examine the relationship between DMGs and DEGs in which DMGs are defined by bases completely methylated in one condition ( $\geq 99\%$ ) and hypomethylated in the other ( $\leq 89\%$ ) (see Fig. S2P to R). This revealed no additional associations, but the association with wild-type bacteria in SPI-2 media replicated once again. We conclude that for the crucial switch between SPI-1 and SPI-2 virulence gene programs, there is no association between m<sup>6</sup>A DNA methylation and transcriptional regulation but that specific mutants may show an association.

**YhdJ plays little role in *Salmonella* physiology under standard conditions important for virulence.** While our data do not support a broad, global correlation between differential methylation and differential expression under most of our tested conditions, particularly comparing wild-type bacteria grown in LB and SPI-2-inducing conditions, this does not rule out that there are discrete examples where methylation and gene expression are causally linked in our data sets. We hypothesized that such instances could be identified by combining our data on methylation and transcriptional patterns under biologically relevant conditions with data from methylase knock-out mutants to reduce our search space to putative sites of regulation. We tested this hypothesis with the YhdJ (the most dynamic methylase in our data set) and Dam (the most well-studied DNA methylase in *Salmonella*) mutants.

RNA-seq on wild-type and  $\Delta yhdJ$  bacteria grown in LB or SPI-2-inducing conditions (GEO [GSE185073](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185073); see Data Set S5) revealed that knocking out *yhdJ* had almost no impact on the transcriptome. Apart from *yhdJ* itself, only 12 genes were differentially expressed in LB (Fig. 6A and Table 2) despite the loss of methylation at all 513 ATGCA\*T sites (see Fig. 3D), and no genes were differentially expressed under SPI-2-inducing conditions (Fig. 6B). Curiously, GO-analysis (79, 80) demonstrated the differentially expressed genes are enriched for *de novo* UMP biosynthetic processes (FDR =  $6.3 \times 10^{-4}$ ) and *de novo* pyrimidine nucleobase biosynthetic processes (FDR =  $8.31 \times 10^{-4}$ ). However, examining these genes further revealed that only two differentially expressed genes contained or were near an ATGCAT sequence (*dppA* and *pyrB*), and only *dppA* was detected to house a methylated ATGCA\*T motif in our replication methylation data set (Table 2), making the mechanism of this differential expression unclear. In agreement with these findings, *yhdJ* deletion had little impact on cellular or virulence phenotypes. The  $\Delta yhdJ$  mutation had no effect on growth in LB or SPI-2-inducing media (Fig. 6C and D), a subtle increase on the amount of observed THP-1 cell infection (Fig. 6E), no effect on replication in THP-1 cells (Fig. 6F), no effect on motility (Fig. 6G), and almost no effect on fitness in intraperitoneal (i.p.) or enteric fever models of mouse infection—though a small increase in the number of  $\Delta yhdJ$  CFU recovered from the spleen relative to the wild type following oral gavage was observed (Table 3). Across all phenotypes, no genetic interaction between *yhdJ* and

#### FIG 4 Legend (Continued)

to be differentially methylated (differentially methylated gene [DMG]) in our binary analysis, differentially expressed (differentially expressed gene [DEG]; FDR < 0.05), differentially methylated and differentially expressed (DM and DE gene), or neither differentially methylated nor expressed. A Fisher exact test was then used to determine whether there was an association between methylation and gene expression. (E to G) Differential methylation is not associated with differential expression. Observed and expected numbers of differentially methylated and differentially expressed genes were not significantly different when comparing uniquely methylated genes in LB versus SPI-2 media (E), wild-type versus  $\Delta metJ$  in LB (F), or wild-type versus  $\Delta metJ$  bacteria in SPI-2 media (G). Uniquely methylated genes are plotted in the condition under which they are methylated (e.g., for panel E, a gene that contains a base that is methylated in LB but not SPI-2 media would be plotted as part of “LB”) but are agnostic to the direction of effect for the expression data. Expected values are calculated by multiplying the frequency of differential methylation by the frequency of differential expression by the total number of genes in the analysis for each condition. Numbers used for the Fisher exact test are shown on the right. Data for panels E and G used data from methylation experiment 1, panel F used the “combined data set.” For panels F and G, the gene *metJ* is removed from the analysis, as it is artificially called both differentially methylated and expressed due to its excision from the genome.



**FIG 5** Quantitative analysis revealed an association between differential methylation and expression between wild-type and  $\Delta metJ$  bacteria. (A) Schematic of RNA-seq and SMRT-seq integration. Each gene in our quantitative analysis was determined to be differentially methylated (differentially methylated gene [DMG]: difference  $\geq 10\%$  methylation across conditions), differentially expressed (differentially expressed gene [DEG]: FDR corrected  $P$  value  $\leq 0.05$ ), differentially methylated and differentially expressed (DM and DE Gene), or neither differentially methylated nor expressed. A Fisher exact test was then used to determine whether there was an association between methylation and gene expression. (B to D) Differential methylation is typically not associated with differential expression. Observed and expected numbers of differentially methylated and differentially expressed genes were not significantly different when comparing uniquely methylated genes in LB versus SPI-2 media (B) or wild-type versus  $\Delta metJ$  bacteria in LB (C); however, a significant enrichment of DEGs was observed in hypermethylated sites in wild-type bacteria grown in SPI-2 media relative to  $\Delta metJ$  (D). Hypermethylated genes are plotted in a condition under which they have increased methylation (e.g., for panel B, a gene that contains a base that is more methylated in LB would be plotted as part of “LB”) but are agnostic to the direction of effect for the expression data. Expected values are calculated by multiplying the frequency of differential methylation by the frequency of differential expression by the total number of genes in the analysis for each condition. Numbers used for the Fisher exact test are shown on the right. Data for panels B and D used data from methylation experiment 1; data for panel C used the “combined data set.” For panels C and D, the gene *metJ* was removed from the analysis, as it is excised from the genome.

*metJ* was detected. Thus, YhdJ has very little impact on transcription or virulence associated phenotypes, and if anything, modestly impairs *Salmonella* virulence.

These findings suggest that YhdJ is almost entirely dispensable under the conditions tested here, despite methylating over 500 sites in the genome and methylation of its ATGCA\**T* motif being the most dynamic under the conditions tested. We questioned whether YhdJ may play roles outside transcription, and whether its dynamic nature is incidental. Supporting this hypothesis, an evolutionary analysis of over 9,000 strains from 10 serovars revealed that the methylase is highly conserved, with the exception of *S. Paratyphi A*, where most strains harbor a 151\* truncation (Fig. 6H). This



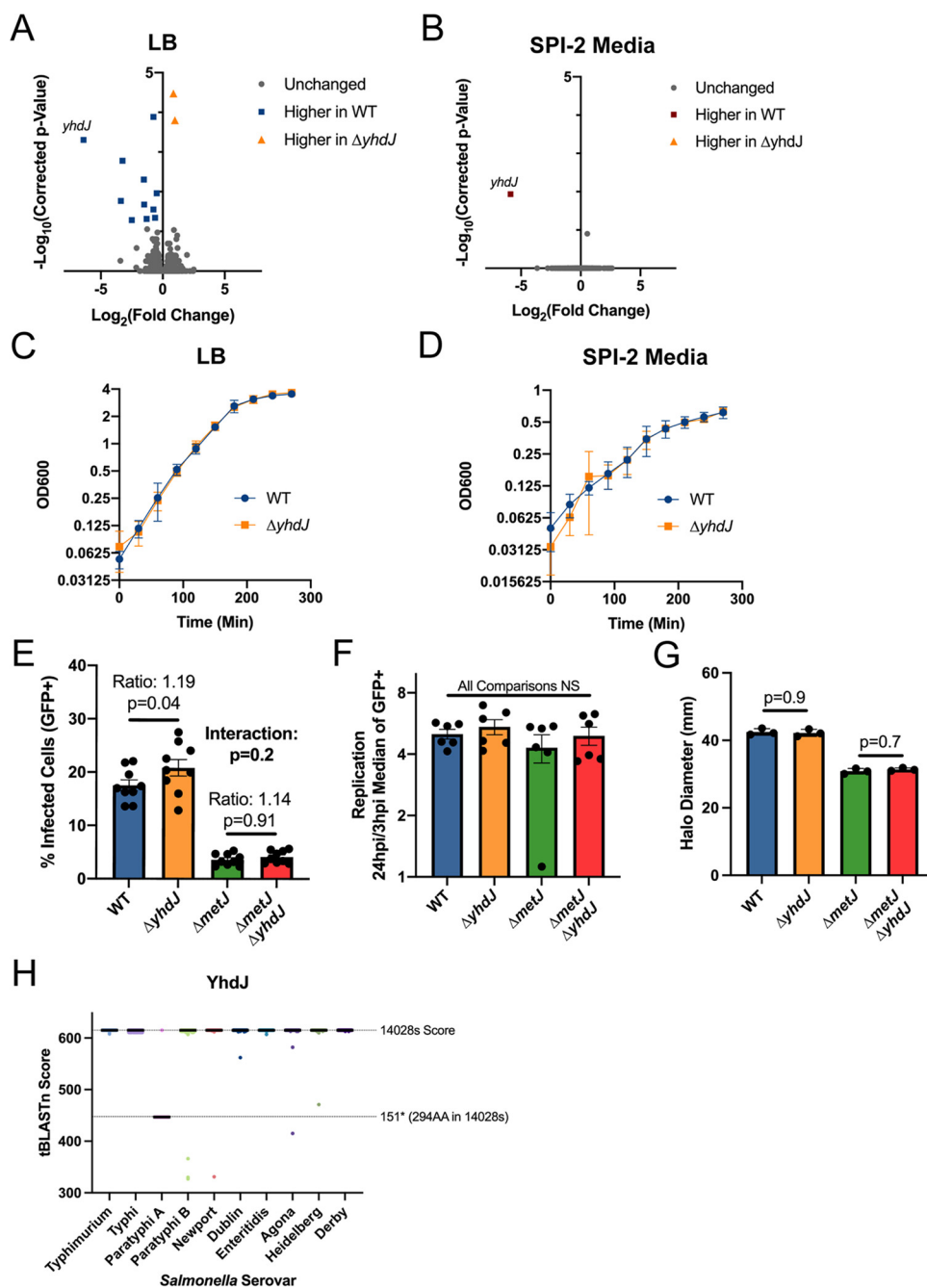
**TABLE 1** Correlation of differential methylation and differential expression in wild-type *S. Typhimurium* grown in SPI-2 inducing media compared to  $\Delta metJ$  bacteria

Genomic location	% methylated (SPI-2)		$\Delta\%$ methylation <sup>a</sup>	STM14 ID	Common name	Relative location	Motif <sup>b</sup>	Log <sub>2</sub> FC <sup>c</sup>
	WT	$\Delta metJ$						
NC_016856.1__1217	99	88	11	STM14_0002	<i>thrA</i>	Genic	GATCA*G	-1.35
NC_016856.1__277289	99	88	11	STM14_0276	<i>ldcC</i>	Genic	GA*TC	0.89
NC_016856.1__288405	98	87	11	STM14_0289	<i>metN</i>	Genic	GA*TC	4.27
NC_016856.1__572813	93	74	19	STM14_0601	<i>sfbB</i>	Genic	CAGA*G	3.25
NC_016856.1__573231	100	90	10	STM14_0601	<i>sfbB</i>	Genic	GA*TC	3.25
NC_016856.1__573265	100	81	19	STM14_0601	<i>sfbB</i>	Genic	GA*TC	3.25
NC_016856.1__573302	100	88	12	STM14_0601	<i>sfbB</i>	Genic	GA*TC	3.25
NC_016856.1__573803	100	90	10	STM14_0602	NA <sup>d</sup>	Genic	GA*TC	3.62
NC_016856.1__861113	100	87	13	STM14_0920	<i>bioB</i>	Genic	GA*TC	-1.22
NC_016856.1__861413	100	88	12	STM14_0920	<i>bioB</i>	Genic	GA*TC	-1.22
NC_016856.1__1048140	98	88	10	STM14_1130	<i>ompF</i>	Genic	GA*TC	1.32
NC_016856.1__1423354	100	89	11	STM14_1619	<i>thrS</i>	Genic	GA*TC	-0.59
NC_016856.1__1733156	76	40	36	STM14_1975	NA	Upstream	Other	1.46
NC_016856.1__1829331	99	86	13	STM14_2086	NA	Genic	GA*TC	3.63
NC_016856.1__1831286	100	71	29	STM14_2087	<i>trpD</i>	Genic	GA*TC	3.71
NC_016856.1__1834122	100	84	16	STM14_2089	<i>trpB</i>	Genic	CAGA*G	2.13
NC_016856.1__2340851	98	81	17	STM14_2702	NA	Genic	GA*TC	2.79
NC_016856.1__2746393	100	88	12	STM14_3133	NA	Genic	GATCA*G	1.80
NC_016856.1__3270917	100	88	12	STM14_3733	<i>metK</i>	Genic	GA*TC	3.21
NC_016856.1__3272719	100	88	12	STM14_3735	NA	Genic	GA*TC	-0.82
NC_016856.1__3722412	100	90	10	STM14_4260	<i>asd</i>	Genic	CAGA*G	2.03
NC_016856.1__4334354	100	87	13	STM14_4936	<i>katG</i>	Genic	GA*TC	5.10
NC_016856.1__4417534	100	90	10	STM14_5030	<i>aceA</i>	Genic	GA*TC	2.42
NC_016856.1__4417670	100	78	22	STM14_5030	<i>aceA</i>	Genic	GA*TC	2.42
NC_016856.1__4423175	100	89	11	STM14_5035	<i>metH</i>	Genic	GA*TC	3.09
NC_016856.1__4424372	100	89	11	STM14_5035	<i>metH</i>	Genic	GA*TC	3.09

<sup>a</sup>That is, WT -  $\Delta metJ$ .<sup>b</sup>\*, The *metJ* gene was excluded from these analyses due to artificial excision from the genome.<sup>c</sup>That is,  $\Delta metJ_{SPI-2}/WT_{SPI-2}$ . FC, fold change.<sup>d</sup>NA, not applicable.

conservation paired with our RNA-seq data suggest that YhdJ methylation is likely important but is unlikely to play a large role in gene regulation.

**Integration of previous  $\Delta dam$  literature with methylomics data reveals *stdA* hypomethylation correlates with expression in LB.** After demonstrating that a  $\Delta yhdJ$  mutant could not be leveraged to identify instances where gene expression and methylation are linked, we turned to the  $\Delta dam$  literature. Previous work on *Salmonella* transcriptomics revealed 17 virulence genes (making up 11 operons) that are differentially expressed between wild-type and  $\Delta dam$  *S. Typhimurium* grown in LB to mid-exponential phase (31). We hypothesized that these sites may show differential methylation and expression between our LB and SPI-2 media, since *Salmonella* deploy radically different virulence programs across these conditions and many of these 17 virulence genes are known to be expressed specifically under only one of these two conditions (14 of 17 were DEGs comparing these two conditions in our data, and the remaining three (STM14\_3654, *stdB*, and *stdC*) lacked high enough expression to analyze). In order to test this hypothesis, we searched for GATC motifs upstream (within 500 bp) of the DEGs. Interestingly, 10/11 genes or operons we examined contain GATC sites within 500 bp upstream of these genes (Table 4). However, only one gene (*stdA*) showed evidence of differential methylation under physiological conditions. Interestingly, *stdA* is the only one of these 11 genes/operons where methylation of its promoter has been extensively studied (58, 81). The three GATC sites had reduced methylation following growth in LB; however, each was only hypomethylated on a single strand (Fig. 7A). Interestingly, this hypomethylation agrees with a previous report that Dam and transcription factors compete for binding to the *stdA* promoter (58), since we also observed significantly higher expression of *stdA* in LB than in SPI-2 media (Fig. 7B). Whether this represents a mechanism by which *stdA* is turned off in SPI-2



**FIG 6** YhdJ has limited impacts on *S. Typhimurium* biology under standard laboratory conditions. (A and B). YhdJ has limited impacts on the *S. Typhimurium* transcriptome in LB (A) and SPI-2-inducing media (B). Corrected *P* values generated by calculating the FDR. (C and D) YhdJ is not required for *S. Typhimurium* growth in LB (C) or SPI-2-inducing media (D). Data from three independent experiments with time points taken every 30 min. Error bars represent the standard errors of the mean. (E and F) YhdJ is not required for *S. Typhimurium* uptake (E) or replication (F) in THP-1 monocytes. Cells were infected for 60 min with *S. Typhimurium* harboring an inducible-GFP plasmid before treatment with gentamicin. GFP was induced for 75 min before analysis by flow cytometry. The percent GFP<sup>+</sup> and the median of the GFP<sup>+</sup> cells were measured 3 and 24 h postinfection. Panel E shows the amount of invasion that occurred by reporting the percentage of infected cells at 3 h postinfection, and panel F shows the replication that occurred after infection by dividing the median of the GFP<sup>+</sup> cells at 24 h postinfection by the median of the GFP<sup>+</sup> cells at 3 h postinfection. Data are from two to three independent experiments; each dot represents an independent replicate, bars mark the mean, and error bars are the standard errors of the mean. For Panel E, data were normalized to the grand mean before plotting or performing statistics, and for panel F statistics were performed on the log transformed values. *P* values were generated by two-way analysis of variance (ANOVA) with Sidak's multiple-comparison test. (G) YhdJ does not impact *S. Typhimurium* motility. Motility on soft  
 (Continued on next page)

**TABLE 2**  $\Delta yhdJ$  differential gene expression in LB medium

Gene ID	Gene name	Log <sub>2</sub> FC <sup>a</sup> ( $\Delta yhdJ$ /WT)	Corrected P value	ATGCAT motif? (relative location)	Methylated?
STM14_4375	<i>dppA</i>	-0.75	2.95E-08	Yes (Genic)	Yes
STM14_4084	<i>yhdJ</i>	-6.39	2.25E-07	No	No
STM14_5353	<i>pyrB</i>	-3.25	1.12E-06	Yes (Genic)	No
STM14_4024	<i>codB</i>	-1.52	4.51E-06	No	No
STM14_0699	<i>cstA</i>	-0.48	1.25E-05	No	No
STM14_5352	<i>pyrI</i>	-3.38	2.32E-05	No	No
STM14_0819	NA	0.85	3.34E-05	No	No
STM14_4495	<i>pyrE</i>	-1.49	3.84E-05	No	No
STM14_5141	<i>acs</i>	-0.74	5.74E-05	No	No
STM14_1558	<i>yeaG</i>	-0.62	0.0001	No	No
STM14_3061	<i>uraA</i>	-1.3	0.0001	No	No
STM14_0078	<i>carB</i>	-2.49	0.0001	No	No
STM14_1885	<i>hdeB</i>	0.99	0.0002	No	No

<sup>a</sup>FC, fold change.

media or is a correlated consequence of increased *stdA* transcription in LB (or vice versa) is unclear, but this finding demonstrates that the phenomenon that García-Pastor et al. describe in  $\Delta dam$  mutants occurs naturally under biologically relevant conditions. For the other 10 virulence genes/operons that are reported to undergo differential gene expression in the *dam* mutant, we found no evidence that differential methylation plays any role during their induction during SPI-1 or SPI-2-inducing conditions.

We also attempted to integrate our data with a recent study that examined genetic heterogeneity in *S. Typhimurium* (82). We examined the 16 hypomethylated sites they identified to determine whether our data set could find the same signatures of hypomethylation. Of the seven sites we were able to find in our data set, four showed signs of hypomethylation (see Table S3). Of these four, two sites upstream of *carA* and *dgoR* showed differential methylation across our conditions; however, no consistent trends were observed with differential gene expression across our two RNA-seq data sets (see Fig. S1G to J). Thus, while we were able to observe an association of gene expression and DNA methylation for one canonical example (*stdA*) during growth in LB versus SPI-2-inducing conditions, we did not find this to be a generalizable phenomenon for other genes reported to be differentially regulated in the *dam* mutant or hypomethylated in *S. Typhimurium*.

**Despite limited changes to the GA\*TC methylome, *metJ* and *dam* interact to influence *S. Typhimurium* invasion and motility.** Having successfully used  $\Delta dam$  transcriptomics to identify one biologically meaningful cooccurrence of differential expression and methylation, we next attempted to leverage  $\Delta dam$  mutants to test our hypothesis that aberrant methylation in  $\Delta metJ$  bacteria contribute to the impact of the  $\Delta metJ$  mutation on invasion and motility (73). Importantly, we observed a small growth defect in  $\Delta dam$  and  $\Delta dam \Delta metJ$  bacteria, and so all bacteria used for these experiments were grown an extra 30 min prior to infection to standardize the growth phase used. Knocking out *dam* only modestly reduced the impact of  $\Delta metJ$  on invasion and is therefore unlikely to be the primary mechanism by which *metJ* deletion impacts

**FIG 6** Legend (Continued)

agar was measured 6 h after inoculating the agar and after migration at 37°C. The data are from three independent experiments; each dot is the average of four to five technical replicates, bars mark the mean, and error bars mark the standard error of the mean. Data were normalized to the grand mean prior to plotting or performing statistics. *P* values were generated by one-way ANOVA with Sidak's multiple-comparison test. (H) YhdJ is conserved across several *Salmonella* serovars. *Salmonella* genomes (1,000 Typhimurium, 1,000 Typhi, 1,000 Paratyphi A, 1,000 Paratyphi B, 999 Newport, 1,000 Dublin, 1,000 Enteritidis, 1,000 Agona, 1,000 Heidelberg, and 79 Derby genomes) were obtained from Enterobase (94, 95). Genomes were combined into a single FASTA file per serovar and blasted against the *S. Typhimurium* strain 14028s YhdJ protein sequence using BLAST+ (96). The BLAST score from the top "n" hits were plotted, where "n" is the number of genomes analyzed for that serovar. A black bar represents the median. Dotted lines represent the BLAST score obtained when blasting the 14028s genome, and the score obtained from the 151\* truncation prevalent in *S. Paratyphi A* serovars.

**TABLE 3** YhdJ does not enhance *S. Typhimurium* fitness in C57BL/6J mice

Parameter <sup>a</sup>	$\Delta metJ/WT$	$\Delta yhdJ/WT$	$\Delta metJ \Delta yhdJ/WT$	$\Delta metJ \Delta yhdJ/\Delta metJ$
i.p. infection (CFU from spleen)				
No. of mice	6	8	7	7
Median (95% CI)	0.23 (0.07–0.47)	1.31 (0.60–1.84)	0.26 (0.13–0.78)	1.09 (0.000065–2.19)
<i>P</i>	0.002	0.3	0.002	0.2
	$\Delta yhdJ/WT$ (spleen)	$\Delta yhdJ/WT$ (ileum)		
Oral infection (CFU from spleen and ileum)				
No. of mice	22	21		
Median (95% CI)	1.65 (0.72–2.22)	1.154 (0.74–1.55)		
<i>P</i>	0.02	0.4		

<sup>a</sup>All mice were age and sex matched, both sexes are represented in all experiments, and all data are from at least two experiments. The median competitive index ("median") was calculated by dividing the number colonies obtained of each genotype at 4 days postinfection. Median > 1 = numerator strain outcompeted denominator. CI, confidence interval. *P* values were calculated using a one-sample *t* test on log-transformed data.

invasion (Fig. 8A). In contrast, impairment in motility caused by *metJ* deletion was completely abrogated in  $\Delta dam$  genetic background, suggesting that *dam* and *metJ* impact motility through the same pathway (Fig. 8B). Importantly, comparing isogenic strains (wild-type versus  $\Delta metJ$ ;  $\Delta dam$  versus  $\Delta dam \Delta metJ$ ;  $\Delta dam pdam$  versus  $\Delta dam \Delta metJ pdam$ ) revealed that complementation of *dam* on a low-copy-number plasmid (pWSK129) restored differences between  $\Delta dam$  and  $\Delta dam \Delta metJ$  bacteria, though we observed that *dam* complementation itself reduces motility further. This is likely due to modest *dam* overexpression, which has previously been reported to be a more potent inhibitor of *S. Typhimurium* 14028s motility than *dam* deletion (34).

**The motility defect of  $\Delta metJ$  *S. Typhimurium* partially depends on *tsr*.** We hypothesized that the genetic interaction between *dam* and *metJ* could signify that differential GA\*TC methylation in the  $\Delta metJ$  mutant suppresses bacterial motility. In striking contrast to this hypothesis, our combined binary data set revealed no genes that were both differentially GA\*TC methylated and expressed (except for the deleted *metJ* itself) (Table 5). We next turned to our percent methylation data to examine whether a shift in methylation could explain differences in flagellar gene regulation between the two bacteria. Comparing percent methylation in both methylation data sets at all GA\*TC methylated sites in which the nearest gene is differentially expressed identified 17 sites that had a  $\geq 10\%$  average difference in methylation between wild-type and  $\Delta metJ$  bacteria (Fig. 8C). Because cooccurrence of differential methylation and differential expression is expected to occur frequently by chance, we sought to limit our analyses to bases most likely to impact gene expression. To do this, we restricted our search to GA\*TC sites that are upstream of differentially expressed genes and found two sites of interest. Specifically, these sites are both strands of a single GATC motif upstream of the chemotaxis gene *tsr* that shows elevated methylation in  $\Delta metJ$  bacteria (Fig. 8D and E and Table 6).

This hypermethylation led us to hypothesize that increased methylation upstream of *tsr* in  $\Delta metJ$  bacteria could decrease *tsr* expression and thereby reduce motility. In line with this hypothesis, replacing *tsr* with a kanamycin resistance cassette partially ablated the ability for the  $\Delta metJ$  mutation to cause a motility defect (Fig. 8F, interaction term *P* = 0.005). Curiously, a search for the methylation-sensitive transcription factor CRP (41) binding motif (AAATGTGATCTAGATCACATTT) in the *tsr* promoter with the MEME FIMO Tool (83) demonstrated that the hypermethylated residue lies within a putative CRP binding site. Together, these data tentatively support a model in which hypermethylation upstream of *tsr* in  $\Delta metJ$  bacteria may contribute to the motility defect. However, additional studies are necessary to confirm a causal relationship.

**Increased methylation in the *flhDC* promoter does not contribute to the  $\Delta metJ$  motility defect.** Given that unlike  $\Delta dam$  (Fig. 8B),  $\Delta tsr$  does not account for the entire impact of *metJ* deletion on motility (Fig. 8F,  $\Delta tsr::kan \Delta metJ/\Delta tsr::kan$  ratio = 0.9), we hypothesized that there may be additional differences in GA\*TC methylation between wild-type and  $\Delta metJ$  bacteria that impact motility. Further examination of our

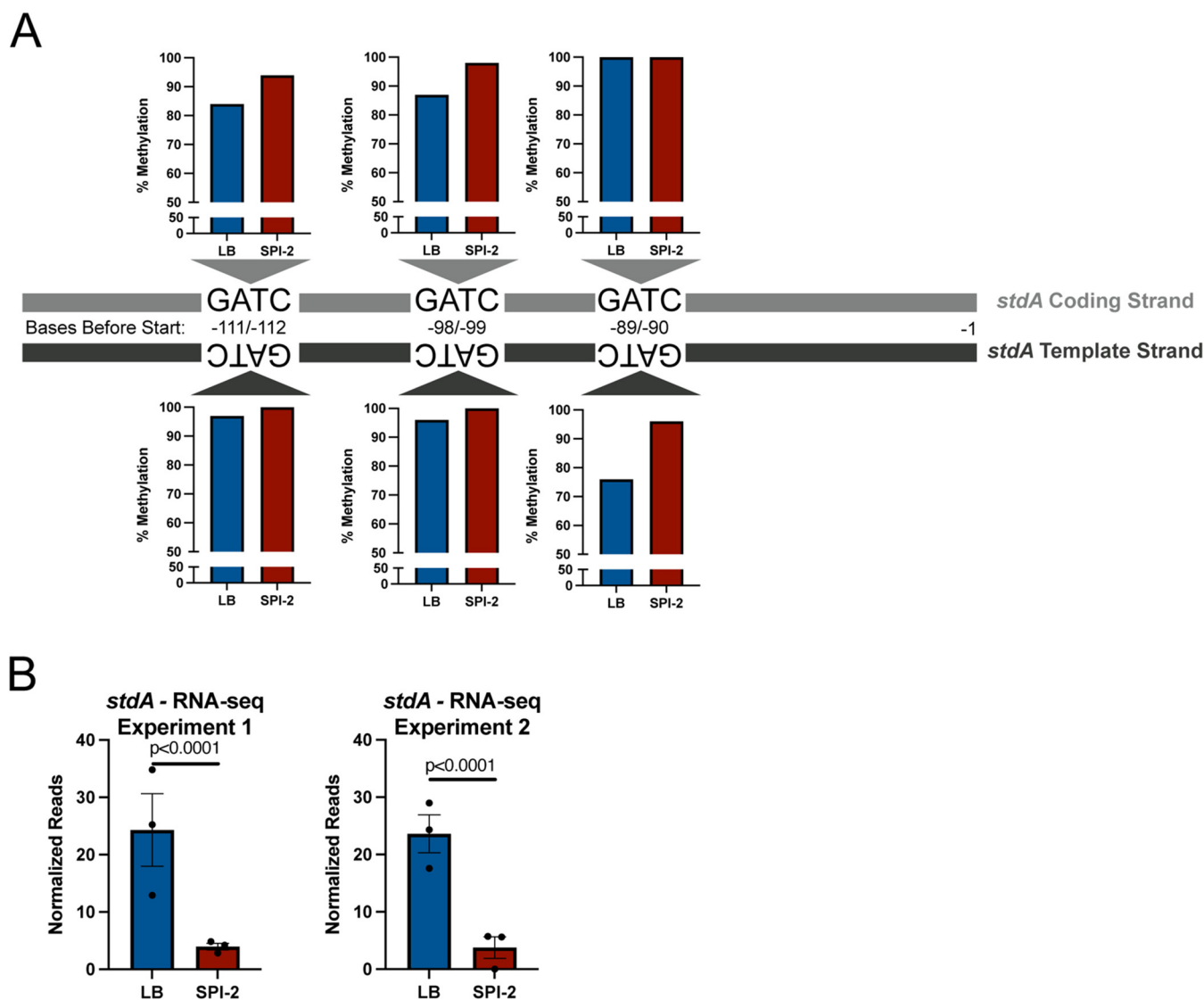
**TABLE 4** Percent methylation of GATC sites in experiment 1 within 500 bp upstream of  $\Delta dam$  differentially expressed virulence genes<sup>a</sup>

Gene/operon (unlisted genes)	Base ID	% methylated in:				
		WT LB	WT SPI-2	$\Delta dam$ LB	$\Delta dam$ SPI-2	
<i>fliC</i>	2060552	100	100	0	0	
	2060553	100	88	0	0	
	2060668	100	100	0	0	
	2060669	100	100	0	0	
	2060836	100	100	0	0	
	2060837	100	100	0	0	
<i>fliD</i> ( <i>fliS-fliT</i> also in operon)	2060552	100	100	0	0	
	2060553	100	88	0	0	
	2060668	100	100	0	0	
	2060669	100	100	0	0	
<i>lppB</i>	1469508	100	100	0	0	
	1469509	93	100	0	0	
	1469697	100	100	0	0	
	1469698	87	80	0	0	
<i>prgH-prgI-prgJ</i> ( <i>prgK</i> also in operon)	None	None	None	None	None	
<i>sicA-sipB-sipC</i>	3051884	92	86	0	0	
	3051885	98	99	0	0	
	3051936	100	100	0	0	
	3051937	100	100	0	0	
	3052008	100	98	0	0	
	3052009	99	100	0	0	
	3052050	100	100	0	0	
	3052051	100	100	0	0	
	<i>sipD</i> ( <i>sipA-iacP</i> also in operon)	3048346	100	100	0	0
		3048347	99	99	0	0
<i>pipC</i> ( <i>sopB</i> also in operon)	1138389	99	90	0	0	
	1138390	96	97	0	0	
	1138429	99	99	0	0	
	1138430	100	100	0	0	
	1138441	99	100	0	0	
	1138442	100	100	0	0	
	1138525	96	100	0	0	
	1138526	98	100	0	0	
<i>cheR</i> ( <i>STM14_2332-STM14_2331-cheZ</i> also in operon)	2026159	90	99	0	0	
	2026160	96	100	0	0	
	2026385	100	100	0	0	
	2026386	100	100	0	0	
	2026415	100	100	0	0	
<i>stdA</i>	2026416	92	100	0	0	
	3211770	76	96	0	0	
	3211771	100	100	0	0	
	3211779	96	100	0	0	
	3211780	87	98	0	0	
	3211792	97	100	0	0	
<i>stdB-stdC</i> ( <i>STM14_3655</i> also in operon)	3211793	84	94	0	0	
	3210924	100	100	0	0	
	3210925	100	100	0	0	
	3211283	99	99	0	0	
<i>STM14_3654-STM14_3653</i>	3211284	89	95	0	0	
	3206809	100	94	0	0	
	3206810	100	100	0	0	

<sup>a</sup>Based on reference 31.

quantitative methylation data set (Table 6) revealed one additional plausible hypothesis: a site in the *flhDC* promoter (−278) that barely missed our 10% threshold (9.5% more methylated in  $\Delta metJ$  bacteria). We decided to test this site as well, since FlhDC make up the master flagellar regulator, and thus modest methylated-mediated regulation of the operon could explain our findings. To test whether differential methylation of the *flhDC* promoter could explain the motility defect in  $\Delta metJ$  bacteria, we performed site directed mutagenesis on the *S. Typhimurium* chromosome to mutate the



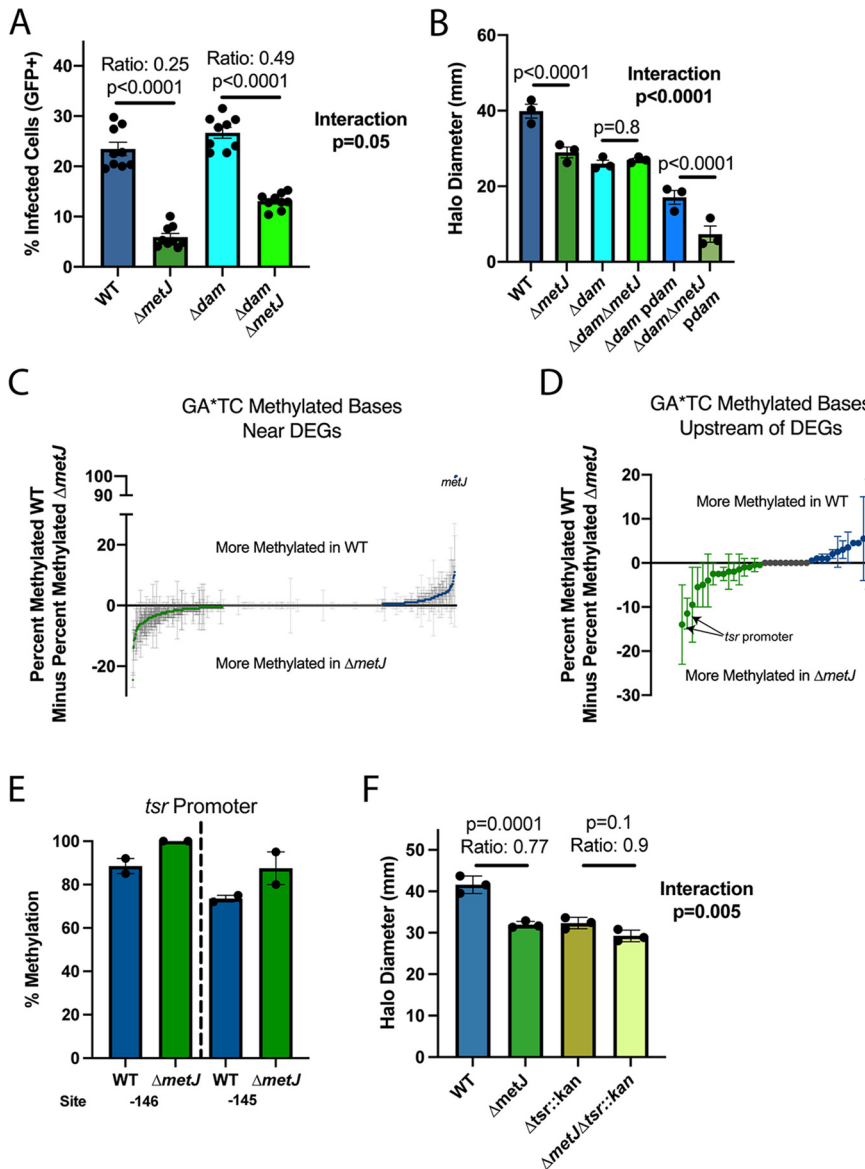


**FIG 7** The *stdA* promoter has differential methylation after growth in LB or SPI-2-inducing media. (A) Schematic of the region upstream of *stdA*. The percent methylation values for each GATC site on both strands are graphed based on data from wild-type bacteria in methylation experiment 1. (B) *stdA* is differentially expressed in wild-type bacteria grown in LB and SPI-2-inducing media. RNA-seq experiment 1 values are from the RNA-seq experiment, including  $\Delta metJ$  bacteria, and are listed in Data Set S4. RNA-seq experiment 2 values are from the experiment, including  $\Delta yhdJ$  bacteria, and are listed in Data Set S5.

base from GATC to GTTC. However, this mutation had no effect on motility in wild-type or  $\Delta metJ$  bacteria (see Fig. S1K), disproving the hypothesis that this site could contribute to the  $\Delta dam$  epistatic effect. Notably, this does not rule out that hypermethylation of this site could play a role in flagellar gene expression in other contexts but does demonstrate that it does not contribute to the  $\Delta metJ$  motility defect.

## DISCUSSION

Here, we demonstrate that at the genome-wide level differential methylation and differential expression are not correlated. Under the critical conditions of SPI-1 or SPI-2 induction, we observed no association between DEGs and DMGs, whether examining binary changes in methylation or quantitative shifts in methylation of  $>10\%$ . However, our results do demonstrate that genome-wide methylation studies of biologically relevant conditions can be integrated with data from methylase knockout mutants to identify methyl-bases that may be coupled with gene expression, as exemplified by the *stdA* and *tsr* examples. Integration of data from future methylomic studies with our



**FIG 8** *dam* is epistatic to  $\Delta metJ$  despite limited changes to the  $\Delta metJ$  GA\*TC methylome. (A) The impacts of  $\Delta metJ$  on invasion partially depend on *dam*. THP-1 monocytes were infected for 60 min with *S. Typhimurium* harboring an inducible-GFP plasmid before treatment with gentamicin. GFP was induced for 75 min before analysis by flow cytometry. The percent GFP<sup>+</sup> was measured at 3 h postinfection. The data are from three experiments; each dot represents an independent replicate, the bars mark the mean, and the error bars are the standard errors of the mean. (B) The impact of  $\Delta metJ$  on motility depends entirely on *dam*. Motility on soft agar was measured 6 h after inoculating the agar and after migration at 37°C. The data are from three independent experiments; each dot is the mean of four to five technical replicates, bars mark the mean, and error bars mark the standard errors of the mean. (C and D) Quantitative analysis reveals subtle changes to the GA\*TC methylome in  $\Delta metJ$  bacteria. Each dot represents the difference average percent methylation of GA\*TC bases in which the closest gene in differentially expressed (C) or GA\*TC bases specifically upstream of differentially expressed genes (D), between WT and  $\Delta metJ$  bacteria grown in LB. The data are in duplicate from the methylation experiment 1 and the replication methylation experiment, with error bars showing the error of the mean. Data from panel D are expanded in Table 4. For panels C and D, any base with greater than or less than 0 differential methylation is colored green (more methylated in  $\Delta metJ$  bacteria) or blue (more methylated in wild-type bacteria). (E) The *tsr* promoter is modestly hypermethylated in  $\Delta metJ$ . Percent methylation is plotted for the -146 and -145 GATC motifs from the methylation experiment 1 and the replication methylation experiment, with error bars showing the errors of the mean. Site numbering is relative to the start codon. (F) The impacts of the  $\Delta metJ$  mutation on motility are partially *tsr* dependent. The data are from three independent experiments; each dot is the mean of four to five technical replicates, bars mark the mean, and error bars mark the standard errors of the mean. For panels A, B, and F, data were normalized to the grand mean prior to plotting or performing statistics and *P* values were generated by two-way ANOVA with Sidak's multiple-comparison test.

**TABLE 5** Unique GA\*TC sites between wild-type and  $\Delta metJ$  *S. Typhimurium* (combined data set)

Genomic location <sup>a</sup>	Base	Genic location	Gene	Product	Differentially expressed?
Unique WT LB GA*TC sites					
None	None	None	None	None	NA <sup>b</sup>
Unique $\Delta metJ$ LB GA*TC sites					
NC_016856.1	1177765	Upstream	STM14_1293	N-Acetylneuraminase Epimerase	No
NC_016856.1	4289018	Downstream	STM14_4888	CDP-diacylglycerol Pyrophosphatase	No

<sup>a</sup>, The *metJ* gene is excluded from these analyses due to artificial excision from the genome.

<sup>b</sup>NA, not applicable.

publicly available data sets could reveal additional naturally occurring instances and potentially important cooccurrence of differential methylation and differential expression. As our work demonstrates, such instances will likely be challenging to identify, since they do not occur more often than expected by chance and therefore do not appear to be a general mechanism of gene regulation in *Salmonella*. In addition, we hope that this work encourages the generation of additional methylomic data sets under diverse and biologically relevant conditions in order to enable more intraspecies comparative methylomics.

A surprising aspect of our work was that the most differentially active methylase we observed, YhdJ, appears to have almost no impact on the *S. Typhimurium* transcriptome under standard conditions. In contrast to Dam which has known impacts on DNA and bacterial replication (10–19), YhdJ also appears to be completely nonessential for *S. Typhimurium* fitness under our growth conditions and in mice. This raises questions about the broader role of DNA methylation, and in particular YhdJ methylation, in the bacterial cell. One tantalizing hypothesis is that YhdJ plays a role in phage defense, which would have been missed studying the conditions here. Alternatively, YhdJ may contribute to physical genomic structural stability under stress conditions, similar to a proposed role for Dam during antibiotic treatment (32). Although these hypotheses could explain why YhdJ does not impact gene expression, they fail to address why we observed reproducible changes in the YhdJ methylome across different conditions. As an answer to this, we speculate that these differences are due to changes in the accessibility of YhdJ to ATGCAT motifs under the different conditions, rather than intentional targeting of YhdJ to these sites. This could be due to differences in other genomic modifications that antagonize YhdJ function, altered protein-DNA interactions that mask ATGCAT sites, and/or changes to the three-dimensional (3D) conformation of the genome that prevent interactions between YhdJ and its motif.

We propose three potential explanations for the lack of a consistent correlation between global m<sup>6</sup>A DNA methylation and gene expression in our data. The first is that while *S. Typhimurium* can and do use m<sup>6</sup>A methylation as a mechanism to promote bistability or otherwise regulate transcription, they do so sparingly. This would suggest that while the canonical examples of this are elegant (12, 21, 22, 29, 30, 36, 54–58, 82), they are rare exceptions to the general rules of *S. Typhimurium* gene regulation. While we are certainly not the first to discover individual sites of differential m<sup>6</sup>A methylation that do not correlate with gene expression (84, 85), this is the first analysis to demonstrate how widespread the phenomenon is in *S. Typhimurium*. The second hypothesis is that three of the four conditions tested here (wild-type or  $\Delta metJ$  bacteria grown in LB or SPI-2-inducing media) are nonrepresentative conditions, whereas our results with the wild-type versus the  $\Delta metJ$  mutant in SPI-2 media are more representative of methylation's relationship with transcription. Notably, while this is possible, these conditions were specifically chosen since they (i) are relevant to the pathogenic capacity of the bacteria, (ii) are relevant to the conditions most frequently studied in laboratory settings, or (iii) disrupt metabolic pathways directly connected to methylation. Therefore, even if methylation plays larger roles in regulating gene expression under other conditions (e.g., nutrient poor conditions at ambient temperature, following phage insult, etc.), our findings would still suggest that most

**TABLE 6** Percent methylation differences for GA\*TC motifs upstream of  $\Delta metJ$  differentially expressed genes following growth in LB medium<sup>a</sup>

Base	Closest gene (STM14 no.)	Closest gene (common name)	% methylation (WT – $\Delta metJ$ )			Gene expression	
			Methylation expt 1	Replication expt	Avg	Log <sub>2</sub> FC ( $\Delta metJ$ /WT)	Corrected <i>P</i>
4802770	STM14_5446	<i>tsr</i>	-23	-5	-14	-1.28	0.006
4802769	STM14_5446	<i>tsr</i>	-15	-8	-11.5	-1.28	0.006
2033749	STM14_2341	<i>flhD</i>	-1	-18	-9.5	-0.53	0.05
3243659	STM14_3699	<i>serA</i>	-10	-1	-5.5	1.48	5.09E-13
4415526	STM14_5029	<i>aceB</i>	-10	0	-5	1.88	6.67E-06
3335887	STM14_3821		2	-10	-4	-1.21	0.04
3400695	STM14_3893		0	-5	-2.5	-1.32	0.007
3400696	STM14_3893		-2	-3	-2.5	-1.32	0.007
4415527	STM14_5029	<i>aceB</i>	-1	-4	-2.5	1.88	6.67E-06
2060553	STM14_2378		2	-6	-2	-1.52	0.0005
3335886	STM14_3821		-4	0	-2	-1.21	0.04
571705	STM14_0600		2	-5	-1.5	4.1	1.63E-39
4415556	STM14_5029	<i>aceB</i>	-3	1	-1	1.88	6.67E-06
4802790	STM14_5446	<i>tsr</i>	-1	-1	-1	-1.28	0.006
3758504	STM14_4305		-2	1	-0.5	-1.41	6.13E-06
3846384	STM14_4392		-1	0	-0.5	3.05	3.03E-07
571704	STM14_0600		0	0	0	4.1	1.63E-39
2060669	STM14_2378		0	0	0	-1.52	0.0005
2746734	STM14_3135	<i>hmpA</i>	0	0	0	1.79	7.28E-05
2746735	STM14_3135	<i>hmpA</i>	0	0	0	1.79	7.27E-05
3243658	STM14_3699	<i>serA</i>	0	0	0	1.48	5.09E-13
3402626	STM14_3894		0	0	0	-1.23	7.87E-08
4185969	STM14_4772		0	0	0	3.17	4.96E-10
4185987	STM14_4772		0	0	0	3.17	4.96E-10
4185988	STM14_4772		0	0	0	3.17	4.96E-10
4415557	STM14_5029	<i>aceB</i>	1	0	0.5	1.88	6.67E-06
1049501	STM14_1130	<i>ompF</i>	1	1	1	1.08	4.46E-07
2060668	STM14_2378		2	0	1	-1.52	0.0005
4185970	STM14_4772		2	0	1	3.17	4.96E-10
3758505	STM14_4305		1	3	2	-1.41	6.13E-06
3846383	STM14_4392		-1	6	2.5	3.05	3.04E-07
2060552	STM14_2378		5	1	3	-1.52	0.0005
4802791	STM14_5446	<i>tsr</i>	7	0	3.5	-1.28	0.006
1049502	STM14_1130	<i>ompF</i>	5	4	4.5	1.08	4.46E-07
3402625	STM14_3894		4	5	4.5	-1.23	7.86E-08
2072703	STM14_2394	<i>fliJ</i>	15	-4	5.5	-1.55	0.006
2072704	STM14_2394	<i>fliJ</i>	19	-6	6.5	-1.55	0.006

<sup>a</sup>, The *metJ* gene was excluded from these analyses due to artificial excision from the genome.

observed *S. Typhimurium* phenomenon are unlikely to be linked to changes in m<sup>6</sup>A methylation. The third possibility is that while m<sup>6</sup>A is the most common modification to the *S. Typhimurium* genome, other modifications (m<sup>5</sup>C, phosphorothioation, etc.) may have more important impacts on gene expression.

In conclusion, through this work we have increased our understanding of the *S. Typhimurium* methylome by defining it as a highly stable system that is largely decoupled from the transcriptome at the genome-wide level. We hope that this work will serve as a reference for how to perform, analyze, and follow-up on DNA methylation studies and that it will help redefine how we think about m<sup>6</sup>A methylation in bacteria.

## MATERIALS AND METHODS

**Bacterial cell culture.** All *Salmonella* strains are derived from *S. Typhimurium* NCTC 12023 (ATCC 14028s) and are included in Table S1. All plasmids are included in Table S1. Chromosomal knockouts were generated by lambda-red recombineering (86). Site-directed mutagenesis of the chromosome was performed using a modified version of lambda-red recombineering, as previously described (87). Complementation plasmids were generated by cut and paste cloning using the pWSK129 plasmid (88). For all experiments, bacteria were maintained on LB (BD, Miller formulation) agar plates, grown in LB media overnight at 37°C at 250 rpm, and subcultured the following morning prior to experiments. The

SPI-2-inducing media is the low phosphate and magnesium (LPM) media from Coombes et al. (59) and contains 5 mM KCl, 7.5 mM  $(\text{NH}_4)_2\text{SO}_4$ , 0.5 mM  $\text{K}_2\text{SO}_4$ , 38 mM glycerol (0.3% [vol/vol]), 0.1% casein hydrolysate, 8  $\mu\text{M}$   $\text{MgCl}_2$ , 337  $\mu\text{M}$   $\text{K}_2\text{HPO}_4$  (pH 5.8), 80 mM MES (pH 5.8), with the final solution pH equal to 5.8. Propagation of temperature-sensitive plasmids occurred at 30°C and were cured at 42°C. Ampicillin was added to media at 100  $\mu\text{g}/\text{mL}$ , kanamycin at 50  $\mu\text{g}/\text{mL}$ , and apramycin at 100  $\mu\text{g}/\text{mL}$ .

**Mammalian cell culture.** THP-1 monocytes from the Duke Cell Culture Facility were cultured at 37°C in 5%  $\text{CO}_2$  in RPMI 1650 media (Invitrogen) supplemented with 10% heat-inactivated fetal bovine serum, 2  $\mu\text{M}$  glutamine, 100 U/mL penicillin-G, and 100 mg/mL streptomycin. Cells used for *Salmonella* gentamicin protection assays were grown in antibiotic-free media 1 h prior to infection.

**Sample preparation for SMRT-Seq.** *S. Typhimurium* were grown overnight, washed once, and subcultured 1:33 in LB for 2 h and 45 min to induce SPI-1 expression or 1:50 in SPI-2-inducing media for 4 h in order to induce SPI-2 expression. A total of  $2 \times 10^9$  bacteria were pelleted, and DNA was extracted using a DNeasy blood and tissue kit (Qiagen). The optional RNase step in the protocol was performed to remove contaminating RNA according to the manufacturer's instructions. DNA was stored at 4°C until library preparation. Multiplexed SMRTbell libraries for sequencing on a PacBio Sequel system were prepared from 1  $\mu\text{g}$  of each microbial gDNA sample. Shearing of gDNA was performed using g-TUBE and centrifugation at  $2,029 \times g$  for 2 min to achieve a target mode size of 10 to 15 kb.

SMRTbell libraries were then prepared using the SMRTbell Express Template Prep kit 2.0. Two pools of eight indexed libraries were prepared. Each pool was then sequenced on a PacBio Sequel SMRTcell using sequencing chemistry 3.0 and 10-h movie length.

**Sample preparation for RNA-Seq.** *S. Typhimurium* were grown overnight, washed once, and subcultured 1:33 in LB for 2 h and 45 min or 1:50 in SPI-2-inducing media for 4 h. A total of  $2 \times 10^9$  bacteria were pelleted at  $5,000 \times g$  for 5 min and resuspended in RNeasy Protect Bacteria lysis reagent (Qiagen) in order to stabilize transcripts. After 5 min, bacteria were repelleted, and resuspended in 200  $\mu\text{L}$  of TE buffer containing lysozyme (15 mg/mL) and 20  $\mu\text{L}$  of proteinase K. Bacteria were vortexed every 2 min for 15 min. Then, 700  $\mu\text{L}$  of  $\beta$ -mercaptoethanol-containing RLT buffer was added. After vortexing, 500  $\mu\text{L}$  of 96% ethanol was added, and the solution was mixed and applied to a RNeasy extraction column (Qiagen). The remainder of the RNeasy protocol was followed according to the manufacturer's instructions. After RNA isolation, 3 to 6  $\mu\text{g}$  of RNA was treated with Turbo DNase (Thermo-Fisher) according to the manufacturer's instructions, with the exception that two successive 30-min DNase treatments were performed. To remove DNase after treatment, the solution was mixed with 350  $\mu\text{L}$  of  $\beta$ -mercaptoethanol-containing RLT buffer, and then 700  $\mu\text{L}$  of 96% ethanol was added. The mixture was then added to a RNeasy MinElute column (Qiagen), and RNA was reisolated according to the manufacturer's instructions.

RNA samples QC was performed with an Agilent fragment analyzer and a Qubit assay on the Perkin-Elmer Victor X2. An Illumina TruSeq Stranded total RNA-Seq kit combined with a Ribo-Zero rRNA removal kit (bacteria) was used to prepare total RNA-seq libraries. Total RNA was first depleted of the rRNA using biotinylated probes that selectively bind to rRNA molecules. The rRNA depleted RNA was then reverse transcribed. During the second-strand synthesis, the cDNA:RNA hybrid is converted into to double-stranded cDNA (dscDNA) and dUTP incorporated into the second cDNA strand, effectively marking the second strand. Illumina sequencing adapters were then ligated to the dscDNA fragments and amplified to produce the final RNA-seq library. The strand marked with dUTP is not amplified, allowing strand specificity sequencing. Libraries were indexed using a dual indexing approach allowing for multiple libraries to be pooled and sequenced on the same sequencing flow cell of an Illumina MiSeq sequencing platform. Before pooling and sequencing, fragment length distribution and library quality were first assessed on a fragment analyzer using a high-sensitivity DNA kit (Agilent Technologies). All libraries were then pooled in equimolar ratio and sequenced. Sequencing was done using 50-bp single-end reads. Once generated, sequence data were demultiplexed and Fastq files generated using Bcl2Fastq conversion software from Illumina.

**SMRT-seq mapping and m<sup>6</sup>A analysis.** m<sup>6</sup>A methylation calls were performed using the pbsmrtpipe base modification and motif detection pipeline (Smrtlink v7.0.1.66975) with *Salmonella enterica* serovar Typhimurium strain 14028s (ASM2216v1) as the reference genome. For sites at or above 50 $\times$  coverage, sites with a phred-based quality score greater than 40 were marked as "1," for strong evidence of methylation; sites with  $\geq 50\times$  coverage but below QV40 were marked as "0," for unlikely to be methylated. For sites below 50 $\times$  coverage, methylation status was not estimated. Assigning methylated bases to motif(s) was performed by comparing the context of the base to known or identified motifs using Microsoft Excel. Motif enrichment was calculated by dividing the frequency of the motif in a given subset (e.g., frequency of the motif in bases only methylated in bacteria grown in LB) and dividing by the frequency of the motif in condition tested (e.g., frequency of the motif among all methylated bases in bacteria grown in LB). Additional methyl-bases were detected on the pWSK29 plasmid harbored in these strains; however, we did not include these bases in our analyses since this plasmid is not involved in the natural lifestyle of *S. Typhimurium*.

**RNA-seq analysis and integration with methylomics.** RNA-seq data were processed using the TrimGalore toolkit ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore)) which employs Cutadapt (89) to trim low-quality bases and Illumina sequencing adapters from the 3' end of the reads. Only reads that were 20 nucleotides or longer after trimming were kept for further analysis. Reads were mapped to the ASM2216v1 version of the *Salmonella enterica* strain 14028S genome and transcriptome (90) using the STAR RNA-seq alignment tool (91). Reads were kept for subsequent analysis if they mapped to a single genomic location. Gene counts were compiled using the HTSeq tool (<http://www-huber.embl.de/users/anders/HTSeq/>). Only genes that had at least 10 reads in any given library were used in subsequent analysis. Normalization and differential expression were carried out using the



DESeq2 (92) Bioconductor (93) package with the R statistical programming environment (<https://www.R-project.org/>). The FDR was calculated to control for multiple hypothesis testing.

Integration of methylomics and RNA-seq analysis occurred in three steps. First, a list of genes present in both analyses was generated. Second, rates of differential expression among (i) the entire list of genes present in both analyses and (ii) differentially methylated genes were called as genes containing 1+ base that was methylated in one condition but not another. Third, expected (frequency of differential expression in the entire list of genes present in both analyses multiplied by the frequency of differential methylation multiplied by the total number of genes in the analysis) and observed differentially methylated and differentially expressed genes were compared. A Fisher exact test was used to determine whether there were statistically significant associations between differential methylation and differential expression.

**Analysis of *yhdJ* across *Salmonella* genomes.** In order to analyze conservation of *yhdJ* across the *Salmonella enterica* genomes, 9,078 genomes (1,000 Typhimurium, 1,000 Typhi, 1,000 Paratyphi A, 1,000 Paratyphi B, 999 Newport, 1,000 Dublin, 1,000 Enteritidis, 1,000 Agona, 1,000 Heidelberg, and 79 Derby genomes) were obtained from the EnteroBase repository (94, 95). Serovars examined here were specifically chosen to test for conservation among a diverse group of *Salmonella*. The specific strains were randomly selected and represented a variety of sources (human, agricultural animal, avian, reptiles, environment, etc.) within serovars, when possible. After downloading the genomes, all genomes of a given serovar were concatenated into a single FASTA file and used for analysis with the BLAST+ command line software (96). The 14028s YhdJ protein sequence was used as query for the pBLASTn program. To determine conservation, the program produced BLAST scores for “*n*” sequences, where *n* = the number of strains tested within each serovar. The BLAST scores were then plotted relative to the BLAST score obtained using the 14028s genome.

**GO-term analysis.** All GO-terms were generated using the Gene Ontology Resource (<http://geneontology.org/>) (79, 80). A PANTHER overrepresentation test was run using the *Salmonella* Typhimurium GO biological process reference, the test used the Fisher exact test, and the correction was based on a calculated false discovery rate. All calculations were run automatically through the web portal software. Any gene that was not present in the GO-term database was “unmapped” and excluded from the analysis.

**Growth curves.** *S. Typhimurium* were grown overnight in LB, subcultured 1:50 into 5 mL of either LB or SPI-2-inducing media, and grown at 37°C at 250RPM. OD600 measurements were taken every 30 min using a spectrophotometer (Pharmacia Biotech Novaspec II).

**Gentamicin protection assay.** Invasion and replication were measured as previously described (97–99). Briefly, bacteria were grown overnight, subcultured 1:33 into 1 mL of LB, and grown for 2 h and 45 min or until all strains entered late-log phase growth ( $OD_{600} = 1.5$  to 2.0) at 37°C with 250 rpm. For any experiment using  $\Delta dam$  bacteria, all bacteria were grown an extra 30 min (3 h and 15 min) so that the  $\Delta dam$  and  $\Delta dam \Delta metJ$  mutants reached late-exponential-phase growth. A total of 100,000 THP-1 monocytes in antibiotic-free media were then infected by *S. Typhimurium* (MOI of 5). At 1 h postinfection, cells were treated with gentamicin (50  $\mu g/mL$ ), and IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) was added 2 h postinfection to induce bacterial GFP expression. At 3 h and 15 min postinfection, cells were read by using a Guava EasyCyte Plus flow cytometer (Millipore). At 22 h and 45 min postinfection, IPTG was added to remaining wells to induce GFP, and at 24 h postinfection the cells were quantified by flow cytometry. The percent host cell invasion was determined by quantifying the number of GFP<sup>+</sup> cells 3 h and 15 min postinfection, and replication was assessed by determining the ratio of the median intensity of GFP<sup>+</sup> cells at 24 h postinfection divided by the median of the GFP<sup>+</sup> cells at 3 h and 15 min postinfection.

**Motility assays.** All strains were cultured overnight in LB, subcultured 1:33 into LB, and grown for 2 h and 45 min at 37°C with 250 rpm. A pipette tip was used to puncture and deliver 2  $\mu L$  of *S. Typhimurium* into the center of a 0.3% LB agar plate. Plates were incubated at 37°C for 6 h before the halo diameter was quantified.

**Murine competitive index experiments.** Mouse studies were approved by the Duke Institutional Animal Care and Use Committee and adhere to the *Guide for the Care and Use of Laboratory Animals* of the National Institutes of Health. All experiments were performed with age- and sex-matched C57BL/6J (7- to 14-week-old) mice. Bacteria were grown overnight, subcultured 1:33, and grown for 2 h and 45 min at 37°C at 250 rpm. The bacteria were then washed and resuspended in phosphate-buffered saline. Inocula were confirmed by plating for CFU. For oral infections, mice were fasted for 12 h before infection and given 100  $\mu L$  of a 10% sodium bicarbonate solution by oral gavage 30 min before infection. Mice then received a 1:1 mixture of two *S. Typhimurium* strains containing either pWSK29 (AmpR) or pWSK129 (KanR) (88), totaling  $10^8$  CFU in 100  $\mu L$ , by oral gavage. For i.p. infections, mice were injected with a 1:1 mixture of two *S. Typhimurium* strains, totaling  $10^3$  CFU in 100  $\mu L$ , into the i.p. space. For both models, tissues were harvested 4 days postinfection, homogenized, and plated on LB agar containing either ampicillin or kanamycin. The competitive index was calculated as follows: (no. of strain A CFU in tissue/no. of strain B CFU in tissue)/(no. of strain A CFU in inoculum/no. of strain B CFU in inoculum). Statistics were calculated by log transforming this ratio from each mouse and comparing to an expected value of 0 using a one-sample *t* test.

**RT-qPCR.** RNA was harvested as described above and used to create cDNA using the iScript cDNA synthesis kit (Bio-Rad Laboratories). qPCR was performed using the iTaq Universal SYBR green Supermix (Bio-Rad Laboratories). Next, 10- $\mu L$  reactions contained 5  $\mu L$  of the supermix, a final 500 nM concentration of each primer, and 2  $\mu L$  of cDNA. Reactions were run on a QuantStudio 3 Thermo Cycler. The cycling conditions were as follows: (i) 95°C for 30 s, (ii) 40 cycles of 95°C for 15 s and 60°C for 60 s, and then (iii) 60°C for 60 s. A melting curve was determined in order to verify single PCR products. The comparative threshold cycle ( $C_T$ ) method was used to quantify transcripts, with the ribosomal *rrs* gene serving as the endogenous control. The fold change represents  $2^{-\Delta\Delta C_T}$ . Oligonucleotides are listed in Table S2.

**Western blotting.** *flhC* was tagged with the 3xFLAG tag using recombineering as previously described (100). *S. Typhimurium* bacteria were grown overnight in LB, subcultured 1:33 in LB at 37°C and 250 rpm until reaching late log phase ( $OD_{600} = 1.5$  to  $2.0$ ), and then pelleted by centrifugation at  $6,000 \times g$  for 5 min. Pellets were resuspended in  $2 \times$  Laemmli buffer (Bio-Rad) with 5% 2-mercaptoethanol and boiled for 10 min, and lysates were run on Mini-Protean TGX stain-free gels (Bio-Rad). After electrophoresis, the gels' total protein dye was activated by a 5-min UV exposure. After transfer onto Immobilon-P low-fluorescence polyvinylidene difluoride membrane (Bio-Rad) using a Hoefer TE77X, blots were probed using an anti-FLAG M2 antibody (Sigma, catalog no. F3165). A fluorescent secondary antibody (LI-COR IRDye) was used to detect bands on a LI-COR Odyssey Classic. Band intensity was quantified using LI-COR Odyssey Imaging System Software v3.0. Total protein was detected by 30 s of UV exposure and quantified using Fiji (101). The graphed relative signal was calculated as follows: (FLAG band intensity/total protein)/(FLAG band intensity in wild-type *flhC:FLAG3* × bacteria/total protein in wild-type *flhC:FLAG3* × bacteria).

**Statistical analyses.** Statistics were performed using GraphPad Prism 9 or Microsoft Excel, except where otherwise noted. Where noted, interexperimental noise was removed from gentamicin protection assays and motility assays prior to data visualization or statistical analysis by standardizing data to the grand mean by multiplying values within an experiment by a constant (the average of all experiments divided by the averages of specific experiments). All statistical tests corresponding to reported *P* values are described in the appropriate figure legends.

**Data availability.** All sequencing data are available in the NCBI's Gene Expression Omnibus (GEO) (102) Super Series ([GSE185077](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185077)). This includes both SMRT-seq experiments ([GSE185578](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185578) and [GSE185501](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185501)), as well as both RNA-seq experiments ([GSE185072](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185072) and [GSE185073](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE185073)). All biological resources are available upon request to Dennis Ko.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, TIF file, 2.8 MB.

**FIG S2**, TIF file, 2 MB.

**TABLE S1**, DOCX file, 0.02 MB.

**TABLE S2**, DOCX file, 0.02 MB.

**TABLE S3**, DOCX file, 0.02 MB.

**DATA SET S1**, XLSX file, 16.2 MB.

**DATA SET S2**, XLSX file, 16.5 MB.

**DATA SET S3**, XLSX file, 6.8 MB.

**DATA SET S4**, XLSX file, 2.6 MB.

**DATA SET S5**, XLSX file, 2.7 MB.

## ACKNOWLEDGMENTS

We thank the Duke University School of Medicine for the use of the Sequencing and Genomic Technologies Shared Resource for performing the library preparations and sequencing experiments referenced throughout this paper. We thank David Corcoran for supervising the genomic analyses performed by J.L.M. and W.C. We thank Kristin Cleveland and Duke DLAR Breeding Core personnel for breeding and maintenance of mouse lines. pREDTKI (Addgene plasmid 51628; <http://n2t.net/addgene:51628>; RRID:Addgene\_51628), pMDIAI (Addgene plasmid 51655; <http://n2t.net/addgene:51655>; RRID:Addgene\_51655), and pKSI-1 (Addgene plasmid 51725; <http://n2t.net/addgene:51725>; RRID:Addgene\_51725) were gifts from Sheng Yang. We also thank all past and present members of the Ko lab, especially Kyle Gibbs, Alejandro Antonia, Alyson Barnes, Rachel Keener, Angela Jones, and Margaret Gaggioli for their helpful discussions about the manuscript. Finally, we are grateful to Stacy Horner and the Duke Molecular Genetics and Microbiology department for use of equipment and shared resources. All schematic images were generated using Biorender.com.

This study was supported by the National Institutes of Health (1F31AI143147 to J.S.B., R01AI118903 to D.C.K., and R21AI144586 to D.C.K.). The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## REFERENCES

1. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* 89:1827–1831. <https://doi.org/10.1073/pnas.89.5.1827>.
2. 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>.

3. Glickman BW, Radman M. 1980. *Escherichia coli* mutator mutants deficient in methylation-instructed DNA mismatch correction. Proc Natl Acad Sci U S A 77:1063–1067. <https://doi.org/10.1073/pnas.77.2.1063>.
4. Pukkila PJ, Peterson J, Herman G, Modrich P, Meselson M. 1983. Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *Escherichia coli*. Genetics 104:571–582. <https://doi.org/10.1093/genetics/104.4.571>.
5. Glickman BW. 1979. Spontaneous mutagenesis in *Escherichia coli* strains lacking 6-methyladenine residues in their DNA: an altered mutational spectrum in dam- mutants. Mutat Res 61:153–162. [https://doi.org/10.1016/0027-5107\(79\)90122-2](https://doi.org/10.1016/0027-5107(79)90122-2).
6. Schlagman SL, Hattman S, Marinus MG. 1986. Direct role of the *Escherichia coli* Dam DNA methyltransferase in methylation-directed mismatch repair. J Bacteriol 165:896–900. <https://doi.org/10.1128/jb.165.3.896-900.1986>.
7. Torreblanca J, Casadesus J. 1996. DNA adenine methylase mutants of *Salmonella* Typhimurium and a novel dam-regulated locus. Genetics 144:15–26. <https://doi.org/10.1093/genetics/144.1.15>.
8. Robinson VL, Oyston PC, Titball RW. 2005. A *dam* mutant of *Yersinia pestis* is attenuated and induces protection against plague. FEMS Microbiol Lett 252:251–256. <https://doi.org/10.1016/j.femsle.2005.09.001>.
9. Watson ME, Jr, Jarisch J, Smith AL. 2004. Inactivation of deoxyadenosine methyltransferase (*dam*) attenuates *Haemophilus influenzae* virulence. Mol Microbiol 53:651–664. <https://doi.org/10.1111/j.1365-2958.2004.04140.x>.
10. Julio SM, Heithoff DM, Provenzano D, Klose KE, Sinsheimer RL, Low DA, Mahan MJ. 2001. DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. Infect Immun 69:7610–7615. <https://doi.org/10.1128/IAI.69.12.7610-7615.2001>.
11. Chao MC, Zhu S, Kimura S, Davis BM, Schadt EE, Fang G, Waldor MK. 2015. A cytosine methyltransferase modulates the cell envelope stress response in the cholera pathogen [corrected]. PLoS Genet 11:e1005666. <https://doi.org/10.1371/journal.pgen.1005666>.
12. Collier J, McAdams HH, Shapiro L. 2007. A DNA methylation ratchet governs progression through a bacterial cell cycle. Proc Natl Acad Sci U S A 104:17111–17116. <https://doi.org/10.1073/pnas.0708112104>.
13. Stephens C, Reisenauer A, Wright R, Shapiro L. 1996. A cell cycle-regulated bacterial DNA methyltransferase is essential for viability. Proc Natl Acad Sci U S A 93:1210–1214. <https://doi.org/10.1073/pnas.93.3.1210>.
14. Zweiger G, Marczynski G, Shapiro L. 1994. A *Caulobacter* DNA methyltransferase that functions only in the predivisive cell. J Mol Biol 235:472–485. <https://doi.org/10.1006/jmbi.1994.1007>.
15. Campbell JL, Kleckner N. 1990. *Escherichia coli* *oriC* and the *dnaA* gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. Cell 62:967–979. [https://doi.org/10.1016/0092-8674\(90\)90271-F](https://doi.org/10.1016/0092-8674(90)90271-F).
16. Campellone KG, Roe AJ, Løbner-Olesen A, Murphy KC, Magoun L, Brady MJ, Donohue-Rolfe A, Tzipori S, Gally DL, Leong JM, Marinus MG. 2007. Increased adherence and actin pedestal formation by dam-deficient enterohaemorrhagic *Escherichia coli* O157:H7. Mol Microbiol 63:1468–1481. <https://doi.org/10.1111/j.1365-2958.2007.05602.x>.
17. Erova TE, Pillai L, Fadl AA, Sha J, Wang S, Galindo CL, Chopra AK. 2006. DNA adenine methyltransferase influences the virulence of *Aeromonas hydrophila*. Infect Immun 74:410–424. <https://doi.org/10.1128/IAI.74.1.410-424.2006>.
18. Mehling JS, Lavender H, Clegg S. 2007. A Dam methylation mutant of *Klebsiella pneumoniae* is partially attenuated. FEMS Microbiol Lett 268:187–193. <https://doi.org/10.1111/j.1574-6968.2006.00581.x>.
19. Taylor VL, Titball RW, Oyston PCF. 2005. Oral immunization with a dam mutant of *Yersinia pseudotuberculosis* protects against plague. Microbiology (Reading) 151:1919–1926. <https://doi.org/10.1099/mic.0.27959-0>.
20. Henderson IR, Owen P. 1999. The major phase-variable outer membrane protein of *Escherichia coli* structurally resembles the immunoglobulin A1 protease class of exported protein and is regulated by a novel mechanism involving Dam and *oxyR*. J Bacteriol 181:2132–2141. <https://doi.org/10.1128/JB.181.7.2132-2141.1999>.
21. Cota I, Blanc-Potard AB, Casadesus J. 2012. STM2209-STM2208 (*opvAB*): a phase variation locus of *Salmonella enterica* involved in control of O-antigen chain length. PLoS One 7:e36863. <https://doi.org/10.1371/journal.pone.0036863>.
22. Broadbent SE, Davies MR, van der Woude MW. 2010. Phase variation controls expression of *Salmonella* lipopolysaccharide modification genes by a DNA methylation-dependent mechanism. Mol Microbiol 77:337–353. <https://doi.org/10.1111/j.1365-2958.2010.07203.x>.
23. Falker S, Schilling J, Schmidt MA, Heussipp G. 2007. Overproduction of DNA adenine methyltransferase alters motility, invasion, and the lipopolysaccharide O-antigen composition of *Yersinia enterocolitica*. Infect Immun 75:4990–4997. <https://doi.org/10.1128/IAI.00457-07>.
24. Sarnacki SH, Marolda CL, Noto Llana M, Giacomodonato MN, Valvano MA, Cerquetti MC. 2009. Dam methylation controls O-antigen chain length in *Salmonella enterica* serovar Enteritidis by regulating the expression of Wzz protein. J Bacteriol 191:6694–6700. <https://doi.org/10.1128/JB.00839-09>.
25. Sarnacki SH, Castañeda M, d R A, Noto Llana M, Giacomodonato MN, Valvano MÁ, Cerquetti MC. 2013. Dam methylation participates in the regulation of PmrA/PmrB and RcsC/RcsD/RcsB two component regulatory systems in *Salmonella enterica* serovar Enteritidis. PLoS One 8:e56474. <https://doi.org/10.1371/journal.pone.0056474>.
26. Murphy KC, Ritchie JM, Waldor MK, Lobner-Olesen A, Marinus MG. 2008. Dam methyltransferase is required for stable lysogeny of the Shiga toxin (Stx2)-encoding bacteriophage 933W of enterohaemorrhagic *Escherichia coli* O157:H7. J Bacteriol 190:438–441. <https://doi.org/10.1128/JB.01373-07>.
27. Gordeeva J, Morozova N, Sierro N, Isaev A, Sinkunas T, Tsvetkova K, Matlashov M, Truncaite L, Morgan RD, Ivanov NV, Siksnys V, Zeng L, Severinov K. 2019. BREX system of *Escherichia coli* distinguishes self from non-self by methylation of a specific DNA site. Nucleic Acids Res 47:253–265. <https://doi.org/10.1093/nar/gky1125>.
28. Torreblanca J, Marques S, Casadesus J. 1999. Synthesis of FinP RNA by plasmids F and pSLT is regulated by DNA adenine methylation. Genetics 152:31–45. <https://doi.org/10.1093/genetics/152.1.31>.
29. Camacho EM, Casadesus J. 2005. Regulation of *traJ* transcription in the *Salmonella* virulence plasmid by strand-specific DNA adenine hemimethylation. Mol Microbiol 57:1700–1718. <https://doi.org/10.1111/j.1365-2958.2005.04788.x>.
30. Blyn LB, Braaten BA, Low DA. 1990. Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. EMBO J 9:4045–4054. <https://doi.org/10.1002/j.1460-2075.1990.tb07626.x>.
31. Balbontin 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>.
32. Cohen NR, Ross CA, Jain S, Shapiro RS, Gutierrez A, Benelky P, Li H, Collins JJ. 2016. A role for the bacterial GATC methylome in antibiotic stress survival. Nat Genet 48:581–586. <https://doi.org/10.1038/ng.3530>.
33. Shell SS, Prestwich EG, Baek SH, Shah RR, Sasseti CM, Dedon PC, Fortune SM. 2013. DNA methylation impacts gene expression and ensures hypoxic survival of *Mycobacterium tuberculosis*. PLoS Pathog 9:e1003419. <https://doi.org/10.1371/journal.ppat.1003419>.
34. Badie G, Heithoff DM, Sinsheimer RL, Mahan MJ. 2007. Altered levels of *Salmonella* DNA adenine methylase are associated with defects in gene expression, motility, flagellar synthesis, and bile resistance in the pathogenic strain 14028 but not in the laboratory strain LT2. J Bacteriol 189:1556–1564. <https://doi.org/10.1128/JB.01580-06>.
35. Oliveira PH, Ribis JW, Garrett EM, Trzilova D, Kim A, Sekulovic O, Mead EA, Pak T, Zhu S, Deikus G, Touchon M, Lewis-Sandari M, Beckford C, Zeitouni NE, Altman DR, Webster E, Oussenko I, Bunyavich S, Aggarwal AK, Bashir A, Patel G, Wallach F, Hamula C, Huprikar S, Schadt EE, Sebra R, van Bakel H, Kasarskis A, Tamayo R, Shen A, Fang G. 2020. Epigenomic characterization of *Clostridioides difficile* finds a conserved DNA methyltransferase that mediates sporulation and pathogenesis. Nat Microbiol 5:166–180. <https://doi.org/10.1038/s41564-019-0613-4>.
36. Brunet YR, Bernard CS, Gavioli M, Lloubes R, Cascales E. 2011. An epigenetic switch involving overlapping fur and DNA methylation optimizes expression of a type VI secretion gene cluster. PLoS Genet 7:e1002205. <https://doi.org/10.1371/journal.pgen.1002205>.
37. Chen L, Paulsen DB, Scruggs DW, Banes MM, Reeks BY, Lawrence ML. 2003. Alteration of DNA adenine methylase (Dam) activity in *Pasteurella multocida* causes increased spontaneous mutation frequency and attenuation in mice. Microbiology (Reading) 149:2283–2290. <https://doi.org/10.1099/mic.0.26251-0>.
38. Heithoff DM, Sinsheimer RL, Low DA, Mahan MJ. 1999. An essential role for DNA adenine methylation in bacterial virulence. Science 284:967–970. <https://doi.org/10.1126/science.284.5416.967>.
39. García-Del Portillo F, Pucciarelli MG, Casadesus J. 1999. DNA adenine methylase mutants of *Salmonella typhimurium* show defects in protein secretion, cell invasion, and M cell cytotoxicity. Proc Natl Acad Sci U S A 96:11578–11583. <https://doi.org/10.1073/pnas.96.20.11578>.



40. Atack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. 2018. Phasevarions of bacterial pathogens: methylomics sheds new light on old enemies. *Trends Microbiol* 26:715–726. <https://doi.org/10.1016/j.tim.2018.01.008>.
41. Sanchez-Romero MA, Casadesus J. 2020. The bacterial epigenome. *Nat Rev Microbiol* 18:7–20. <https://doi.org/10.1038/s41579-019-0286-2>.
42. Beaulaurier J, Schadt EE, Fang G. 2019. Deciphering bacterial epigenomes using modern sequencing technologies. *Nat Rev Genet* 20:157–172. <https://doi.org/10.1038/s41576-018-0081-3>.
43. Marinus MG, Lobner-Olesen A. 2014. DNA methylation. *EcoSal Plus* 6. <https://doi.org/10.1128/ecosalplus.ESP-0003-2013>.
44. 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>.
45. Rand AC, Jain M, Eizenga JM, Musselman-Brown A, Olsen HE, Akeson M, Paten B. 2017. Mapping DNA methylation with high-throughput nanopore sequencing. *Nat Methods* 14:411–413. <https://doi.org/10.1038/nmeth.4189>.
46. Tourancheau A, Mead EA, Zhang XS, Fang G. 2021. Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nat Methods* 18:491–498. <https://doi.org/10.1038/s41592-021-01109-3>.
47. 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>.
48. Cota I, Bunk B, Sproer C, Overmann J, Konig C, Casadesus J. 2016. OxyR-dependent formation of DNA methylation patterns in OvpABOFF and OvpABON cell lineages of *Salmonella enterica*. *Nucleic Acids Res* 44:3595–3609. <https://doi.org/10.1093/nar/gkv1483>.
49. Erill I, Puigvert M, Legrand L, Guarischi-Sousa R, Vandecasteele C, Setubal JC, Genin S, Guidot A, Valls M. 2017. Comparative analysis of *Ralstonia solanacearum* methylomes. *Front Plant Sci* 8:504.
50. Modlin SJ, Conkle-Gutierrez D, Kim C, Mitchell SN, Morrissey C, Weinrick BC, Jacobs WR, Ramirez-Busby SM, Hoffner SE, Valafar F. 2020. Drivers and sites of diversity in the DNA adenine methylomes of 93 *Mycobacterium tuberculosis* complex clinical isolates. *Elife* 9. <https://doi.org/10.7554/eLife.58542>.
51. Kahramanoglou C, Prieto AI, Khedkar S, Haase B, Gupta A, Benes V, Fraser GM, Luscombe NM, Seshasayee AS. 2012. Genomics of DNA cytosine methylation in *Escherichia coli* reveals its role in stationary phase transcription. *Nat Commun* 3:886. <https://doi.org/10.1038/ncomms1878>.
52. Davis-Richardson AG, Russell JT, Dias R, McKinlay AJ, Canepa R, Fagen JR, Rusoff KT, Drew JC, Kolaczowski B, Emerich DW. 2016. Integrating DNA methylation and gene expression data in the development of the soybean-bradyrhizobium N<sub>2</sub>-fixing symbiosis. *Front Microbiol* 7:518.
53. Srivastava A, Murugaiyan J, Garcia JAL, De Corte D, Hoetzinger M, Eravci M, Weise C, Kumar Y, Roesler U, Hahn MW, Grossart H-P. 2020. Combined methylome, transcriptome and proteome analyses document rapid acclimatization of a bacterium to environmental changes. *Front Microbiol* 11:544785. <https://doi.org/10.3389/fmicb.2020.544785>.
54. Peterson SN, Reich NO. 2008. Competitive Lrp and Dam assembly at the pap regulatory region: implications for mechanisms of epigenetic regulation. *J Mol Biol* 383:92–105. <https://doi.org/10.1016/j.jmb.2008.07.086>.
55. Hernday A, Krabbe M, Braaten B, Low D. 2002. Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci U S A* 99(Suppl 4):16470–16476. <https://doi.org/10.1073/pnas.182427199>.
56. Waldron DE, Owen P, Dorman CJ. 2002. Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in *Escherichia coli*. *Mol Microbiol* 44:509–520. <https://doi.org/10.1046/j.1365-2958.2002.02905.x>.
57. Wallecha A, Munster V, Correnti J, Chan T, van der Woude M. 2002. Dam- and OxyR-dependent phase variation of *agn43*: essential elements and evidence for a new role of DNA methylation. *J Bacteriol* 184:3338–3347. <https://doi.org/10.1128/JB.184.12.3338-3347.2002>.
58. Garcia-Pastor L, Sanchez-Romero MA, Jakomin M, Puerta-Fernandez E, Casadesus J. 2019. Regulation of bistability in the std fimbrial operon of *Salmonella enterica* by DNA adenine methylation and transcription factors HdfR, StdE, and StdF. *Nucleic Acids Res* 47:7929–7941. <https://doi.org/10.1093/nar/gkz530>.
59. Coombes BK, Brown NF, Valdez Y, Brumell JH, Finlay BB. 2004. Expression and secretion of *Salmonella* pathogenicity island-2 virulence genes in response to acidification exhibit differential requirements of a functional type III secretion apparatus and SsaL. *J Biol Chem* 279:49804–49815. <https://doi.org/10.1074/jbc.M404299200>.
60. Galan JE, Curtiss R, III. 1989. Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. *Proc Natl Acad Sci U S A* 86:6383–6387. <https://doi.org/10.1073/pnas.86.16.6383>.
61. Shea JE, Hensel M, Gleeson C, Holden DW. 1996. Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* 93:2593–2597. <https://doi.org/10.1073/pnas.93.6.2593>.
62. Ochman H, Soncini FC, Solomon F, Groisman EA. 1996. Identification of a pathogenicity island required for *Salmonella* survival in host cells. *Proc Natl Acad Sci U S A* 93:7800–7804. <https://doi.org/10.1073/pnas.93.15.7800>.
63. Hondorp ER, Matthews RG. 2006. Methionine. *EcoSal Plus* 2. <https://doi.org/10.1128/ecosalplus.3.6.1.7>.
64. Cantoni GL. 1953. S-Adenosylmethionine; a new intermediate formed enzymatically from L-methionine and adenosinetriphosphate. *J Biol Chem* 204:403–416. [https://doi.org/10.1016/S0021-9258\(18\)66148-4](https://doi.org/10.1016/S0021-9258(18)66148-4).
65. Bigaud E, Corrales FJ. 2016. Methylthioadenosine (MTA) regulates liver cells proteome and methylproteome: implications in liver biology and disease. *Mol Cell Proteomics* 15:1498–1510. <https://doi.org/10.1074/mcp.M115.055772>.
66. Pascale RM, Simile MM, Satta G, Seddaiu MA, Daino L, Pinna G, Vinci MA, Gaspa L, Feo F. 1991. Comparative effects of L-methionine, S-adenosyl-L-methionine and 5'-methylthioadenosine on the growth of preneoplastic lesions and DNA methylation in rat liver during the early stages of hepatocarcinogenesis. *Anticancer Res* 11:1617–1624.
67. Woodcock DM, Adams JK, Allan RG, Cooper IA. 1983. Effect of several inhibitors of enzymatic DNA methylation on the *in vivo* methylation of different classes of DNA sequences in a cultured human cell line. *Nucleic Acids Res* 11:489–499. <https://doi.org/10.1093/nar/11.2.489>.
68. Yi P, Melnyk S, Pogribna M, Pogribny IP, Hine RJ, James SJ. 2000. Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. *J Biol Chem* 275:29318–29323. <https://doi.org/10.1074/jbc.M002725200>.
69. Hoffman DR, Cornatzer WE, Duerre JA. 1979. Relationship between tissue levels of S-adenosylmethionine, S-adenylhomocysteine, and transmethylation reactions. *Can J Biochem* 57:56–65. <https://doi.org/10.1139/o79-007>.
70. Zappia V, Zydek-Cwick R, Schlenk F. 1969. The specificity of S-adenosylmethionine derivatives in methyl transfer reactions. *J Biol Chem* 244:4499–4509. [https://doi.org/10.1016/S0021-9258\(18\)94346-2](https://doi.org/10.1016/S0021-9258(18)94346-2).
71. Mull L, Ebbs ML, Bender J. 2006. A histone methylation-dependent DNA methylation pathway is uniquely impaired by deficiency in *Arabidopsis* S-adenosylhomocysteine hydrolase. *Genetics* 174:1161–1171. <https://doi.org/10.1534/genetics.106.063974>.
72. Barber JR, Clarke S. 1984. Inhibition of protein carboxyl methylation by S-adenosyl-L-homocysteine in intact erythrocytes: physiological consequences. *J Biol Chem* 259:7115–7122. [https://doi.org/10.1016/S0021-9258\(17\)39845-9](https://doi.org/10.1016/S0021-9258(17)39845-9).
73. Bourgeois JS, Zhou D, Thurston TLM, Gilchrist JJ, Ko DC. 2018. Methylthioadenosine suppresses salmonella virulence. *Infect Immun* 86:e00429-18. <https://doi.org/10.1128/IAI.00429-18>.
74. Pirone-Davies C, Hoffmann M, Roberts RJ, Muruvanda T, Timme RE, Strain E, Luo Y, Payne J, Luong K, Song Y, Tsai Y-C, Boitano M, Clark TA, Korlach J, Evans PS, Allard MW. 2015. Genome-wide methylation patterns in *Salmonella enterica* subsp. *enterica* serovars. *PLoS One* 10:e0123639. <https://doi.org/10.1371/journal.pone.0123639>.
75. Broadbent SE, Balbontin R, Casadesus J, Marinus MG, van der Woude M. 2007. YhdJ, a nonessential CcrM-like DNA methyltransferase of *Escherichia coli* and *Salmonella enterica*. *J Bacteriol* 189:4325–4327. <https://doi.org/10.1128/JB.01854-06>.
76. Sanchez-Romero MA, Olivenza DR, Gutierrez G, Casadesus J. 2020. Contribution of DNA adenine methylation to gene expression heterogeneity in *Salmonella enterica*. *Nucleic Acids Res* 48:11857–11867. <https://doi.org/10.1093/nar/gkaa730>.
77. Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. *Nucleic Acids Res* 43:W39–W49. <https://doi.org/10.1093/nar/gkv416>.
78. Kröger C, Colgan A, Srikanth S, Händler K, Sivasankaran SK, Hammarlöf DL, Canals R, Grissom JE, Conway T, Hokamp K, Hinton JCD. 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. *Cell Host Microbe* 14:683–695. <https://doi.org/10.1016/j.chom.2013.11.010>.
79. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. *Nat Genet* 25:25–29. <https://doi.org/10.1038/75556>.

80. Gene Ontology Consortium. 2021. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res* 49(D1):D325–D334. <https://doi.org/10.1093/nar/gkaa1113>.
81. Jakomin M, Chessa D, Baumler AJ, Casadesus J. 2008. Regulation of the *Salmonella enterica* *std* fimbrial operon by DNA adenine methylation. *SeqA*, and *HdfR*. *J Bacteriol* 190:7406–7413. <https://doi.org/10.1128/JB.01136-08>.
82. Sanchez-Romero MA, Olivenza DR, Gutierrez G, Casadesus J. 2020. Contribution of DNA adenine methylation to gene expression heterogeneity in *Salmonella enterica*. *Nucleic Acids Res* 48:11857–11867. <https://doi.org/10.1093/nar/gkaa730>.
83. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27:1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>.
84. van der Woude M, Hale WB, Low DA. 1998. Formation of DNA methylation patterns: nonmethylated GATC sequences in gut and pap operons. *J Bacteriol* 180:5913–5920. <https://doi.org/10.1128/JB.180.22.5913-5920.1998>.
85. Lewis EB, Chen E, Culyba MJ. 2022. DNA cytosine methylation at the *lexA* promoter of *Escherichia coli* is stationary phase specific. *G3 (Bethesda)* 12:jkab409. <https://doi.org/10.1093/g3journal/jkab409>.
86. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97:6640–6645. <https://doi.org/10.1073/pnas.120163297>.
87. Yang J, Sun B, Huang H, Jiang Y, Diao L, Chen B, Xu C, Wang X, Liu J, Jiang W, Yang S. 2014. High-efficiency scarless genetic modification in *Escherichia coli* by using lambda red recombination and I-SceI cleavage. *Appl Environ Microbiol* 80:3826–3834. <https://doi.org/10.1128/AEM.00313-14>.
88. Wang RF, Kushner SR. 1991. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* 100:195–199. [https://doi.org/10.1016/0378-1119\(91\)90366-J](https://doi.org/10.1016/0378-1119(91)90366-J).
89. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10–12. <https://doi.org/10.14806/ej.17.1.200>.
90. Kersey PJ, Staines DM, Lawson D, Kulesha E, Derwent P, Humphrey JC, Hughes DST, Keenan S, Kerhornou A, Koscielny G, Langridge N, McDowall MD, Megy K, Maheswari U, Nuhn M, Paulini M, Pedro H, Toneva I, Wilson D, Yates A, Birney E. 2012. ENSEMBL genomes: an integrative resource for genome-scale data from non-vertebrate species. *Nucleic Acids Res* 40:D91–97. <https://doi.org/10.1093/nar/gkr895>.
91. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
92. 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>.
93. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MacDonald J, Obenchain V, Oleś AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. 2015. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* 12:115–121. <https://doi.org/10.1038/nmeth.3252>.
94. Alikhan NF, Zhou Z, Sergeant MJ, Achtman M. 2018. A genomic overview of the population structure of *Salmonella*. *PLoS Genet* 14:e1007261. <https://doi.org/10.1371/journal.pgen.1007261>.
95. Achtman M, Zhou Z, Alikhan N-F, Tyne W, Parkhill J, Cormican M, Chiou C-S, Torpdahl M, Litrup E, Prendergast DM, Moore JE, Strain S, Kornschöber C, Meinersmann R, Uesbeck A, Weill F-X, Coffey A, Andrews-Polymenis H, Curtiss Rd R, Fanning S. 2020. Genomic diversity of *Salmonella enterica*: the UoWUCC 10K genomes project. *Wellcome Open Res* 5:223. <https://doi.org/10.12688/wellcomeopenres.16291.1>.
96. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
97. Alvarez MI, Glover LC, Luo P, Wang L, Theusch E, Oehlers SH, Walton EM, Tram TTB, Kuang Y-L, Rotter JI, McClean CM, Chinh NT, Medina MW, Tobin DM, Dunstan SJ, Ko DC. 2017. Human genetic variation in VAC14 regulates *Salmonella* invasion and typhoid fever through modulation of cholesterol. *Proc Natl Acad Sci U S A* 114:E7746–E7755. <https://doi.org/10.1073/pnas.1706070114>.
98. Jaslow SL, Gibbs KD, Fricke WF, Wang L, Pittman KJ, Mammel MK, Thaden JT, Fowler VG, Hammer GE, Effenbein JR, Ko DC. 2018. *Salmonella* activation of STAT3 signaling by SarA effector promotes intracellular replication and production of IL-10. *Cell Rep* 23:3525–3536. <https://doi.org/10.1016/j.celrep.2018.05.072>.
99. Wang L, Pittman KJ, Barker JR, Salinas RE, Stanaway IB, Williams GD, Carroll RJ, Balmat T, Ingham A, Gopalakrishnan AM, Gibbs KD, Antonia AL, Heitman J, Lee SC, Jarvik GP, Denny JC, Horner SM, DeLong MR, Valdivia RH, Crosslin DR, Ko DC, eMERGE Network. 2018. An atlas of genetic variation linking pathogen-induced cellular traits to human disease. *Cell Host Microbe* 24:308–323. <https://doi.org/10.1016/j.chom.2018.07.007>.
100. Uzzau S, Figueroa-Bossi N, Rubino S, Bossi L. 2001. Epitope tagging of chromosomal genes in *Salmonella*. *Proc Natl Acad Sci U S A* 98:15264–15269. <https://doi.org/10.1073/pnas.261348198>.
101. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9:676–682. <https://doi.org/10.1038/nmeth.2019>.
102. Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30:207–210. <https://doi.org/10.1093/nar/30.1.207>.