bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Integration of the *Salmonella* Typhimurium methylome and transcriptome following environmental or metabolic perturbation reveals DNA methylation and transcriptional regulation are largely decoupled

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

Affiliations

[a] Department of Molecular Genetics and Microbiology, School of Medicine, Duke University, Durham, NC, 27710, USA
[b] University Program in Genetics and Genomics, Duke University, Durham, NC, 27710, USA
[c] Center for Genomics and Computational Biology, Duke University, Durham, NC, 27710, USA.
[d] Division of Infectious Diseases, Department of Medicine, School of Medicine, Duke University, Durham, NC, 27710, USA

*To whom correspondence should be addressed: Dennis C. Ko, 0048B CARL Building Box

3053, 213 Research Drive, Durham, NC 27710. 919-684-5834. dennis.ko@duke.edu.

@denniskoHiHOST

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Despite being in a golden age of prokaryotic epigenomics, little work has systematically examined the plasticity and functional impacts of the bacterial DNA methylome. Here, we leveraged SMRT sequencing to examine the m^6A DNA methylome of two *Salmonella enterica* ser. Typhimurium strains: 14028s and a Δ*metJ* mutant with derepressed methionine metabolism, grown in Luria Broth or a media that simulates the intracellular environment. We find that the methylome is remarkably static—over 95% of adenosine bases retain their methylation status across conditions. Integration of methylation with transcriptomic data revealed no correlation between methylation and gene expression. Further, examining the transcriptome in Δ*yhdJ* bacteria, lacking the m^6A methylase with the most dynamic methylation pattern in our dataset, revealed little evidence of YhdJ-mediated gene regulation. Curiously, despite G(m^6A)TC motifs being particularly resistant to change across conditions, we found that the Dam methylase is required for the Δ*metJ* motility defect. This Δ*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.

Introduction

Until recently, systematically understanding how the prokaryotic DNA methylome affects bacterial physiology has been an unachievable task. Unlike eukaryotes where m^5C DNA methylation is highly abundant and can be detected with bisulfate sequencing (1), prokaryotic genomes primarily house m^6A 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 both in the contexts of restriction-modification systems (reviewed (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), fimbriae 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 while restriction-modification systems targeted foreign DNA, recent work on "phasevarions" have found restriction-modification systems can indeed have dramatic impacts on the genome (reviewed (40)). A more complete history of the associations between methylases and phenotypes can be found in recent reviews (41-43).

While early epigenome studies are valuable for the insights they provide, their dependence on low-throughput and relatively blunt approaches (*e.g.*, restriction enzyme digests paired with southern blotting to infer methylation) meant that there were technical hurdles that limited mechanistic understanding of DNA methylation. In particular, these approaches could not be leveraged to address if and how genome-wide changes in DNA methylation associate with changes to cellular processes. However, the discovery that sequencing data from the Pac-Bio SMRT-sequencing (44) and Oxford Nanopore sequencing (45,46) systems can be repurposed to detect m^6A has heralded a golden age of prokaryotic DNA methylomics. These technological breakthroughs were rapidly applied to cataloging prokaryotic 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^6A)TC patterns within in the *opvAB* promoter in response to phage insult

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(48), building on phenotypic observations made a decade earlier (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 (11,32,51,52). 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 datasets. Further, this approach is particularly problematic as the very basis of methylation-mediated gene regulation depends on the methylome being flexible, but how much methylation changes is unclear. For instance, studies looking at *Vibrio cholera* growth stages (11) and *E. coli* stress response (32) have found the methylome to be highly resistant to change, while work with *Bradyrhizobium diazoefficiens* has found modest methylome changes during differentiation (52). Therefore, methylome plasticity, and by extension the role the methylome can play in gene regulation in nature, likely depends on the specific bacterial species, methylases, and conditions examined.

An additional shortcoming of many methylation studies is that methylation sites in promoters are often reported as evidence of methylation-mediated regulation, without 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* (30,53,54), *opvAB* (21), *agn43* (55,56), *gtr* (22), the *std* operon (57), *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. Even the example above in which methylation patterns in in the *opvAB* promoter were found to be phage regulated does not meet this criteria, as methylation at these sites had already been identified by traditional methods to impact transcription (21). 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.

In this paper, we perform a series of SMRT-seq and RNA-seq experiments to understand the role of DNA methylation in regulating *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) gene expression under different environmental conditions. Specifically, we studied conditions that activate the motility and *Salmonella* Pathogenicity Island-1 (SPI-1) pathways (growth in LB until late-log phase) and conditions that activate the *Salmonella* Pathogenicity Island-2 (SPI-2) pathway (growth in LPM media (58) until late-log phase). As methionine metabolism is intimately connected to methylation, we also examined the changes in methylation associated with derepressed methionine metabolism using a Δ*metJ* mutant. In general, we find that DNA methylation is extremely stable across conditions and is broadly decoupled from gene expression. Our findings redefine the *S.* Typhimurium epigenome and support caution when interpreting methylation datasets.

Materials and Methods

Bacterial cell culture

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

All *Salmonella* strains are derived from *S*. Typhimurium NCTC 12023 (ATCC 14028s) and are included in **Supplemental Table 1**. All plasmids are included in **Supplemental Table 2**. Chromosomal knockouts were generated by lambda-red recombineering (59). Site-directed mutagenesis of the chromosome was performed using a modified version of lambda-red recombineering, as previously described (60). Complementation plasmids were generated by cut and paste cloning using the pWSK129 plasmid (61). 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*. (58) and contains 5 mM KCl, 7.5mM $(NH_4)_2SO_4$, 0.5mM K_2SO_4, 38mM glycerol (0.3% volume/volume), 0.1% casein hydrolysate, 8μM $MgCl_2$, 337μM K_2HPO_4 (pH 5.8), 80mM 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 100μg/mL, kanamycin at 50μg/mL, and apramycin at 100μg/mL.

Mammalian cell culture

THP-1 monocytes from the Duke Cell Culture Facility were cultured at 37°C in 5% CO2 in RPMI 1650 media (Invitrogen) supplemented with 10% heat-inactivated FBS, 2μM glutamine, 100 U/mL penicillin-G, and 100mg/mL streptomycin. Cells used for *Salmonella* gentamicin protection assays were grown in antibiotic free media one hour prior to infection.

Sample preparation for SMRT-Seq

S. Typhimurium were grown overnight, washed once, and subcultured 1:33 in LB for 2 hours and 45 minutes to induce SPI-1 expression, or 1:50 in SPI-2 inducing media for 4 hours in order to induce SPI-2 expression. $2x10^9$ 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, per manufacturer instructions. DNA was stored at 4˚C until library preparation. Multiplexed SMRTbell libraries for sequencing on the PacBio Sequel system were prepare from 1 μg of each microbial gDNA sample. Shearing of gDNA was performed using g-TUBE and centrifugation at 2029 x g for 2 minutes to achieve a target mode size of 10 kb - 15 kb. SMRTbell libraries were then prepared using the SMRTbell Express Template Prep Kit 2.0. Two pools of 8 indexed libraries were prepared. Each pool was then sequenced on a PacBio Sequel SMRTcell using sequencing chemistry 3.0 and 10 hour movie length.

Sample preparation for RNA-Seq

S. Typhimurium were grown overnight, washed once, and subcultured 1:33 in LB for 2 hours and 45 minutes, or 1:50 in SPI-2 inducing media for 4 hours. $2x10^9$ bacteria were pelleted at 5,000xg for 5 minutes, and resuspended in RNAprotect Bacteria Reagent (Qiagen) in order to stabilize transcripts. After 5 minutes, bacteria were re-pelleted, and resuspended in 200μL of TE Buffer containing lysozyme (15mg/mL) and 20μL of Proteinase K. Bacteria were vortexed every two minutes for 15 minutes. 700μL of β-mercaptoethanol-containing RLT buffer was added. After vortexing, 500μL 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 per manufacturer instructions. After RNA isolation, 3-6μg of RNA was treated with Turbo DNase (Thermo-Fisher) per manufacturer instructions, with the exception that two successive 30-minute DNase treatments were performed. To remove DNase after treatment, the solution was mixed with 350 μL of β-

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

mercaptoethanol-containing RLT buffer, and then 700µL of 96% ethanol was added. The mixture was then added to a RNeasy MinElute column (Qiagen) and RNA was reisolated according to manufacturer instructions.

RNA samples QC was performed with an Agilent Fragment Analyzer and a Qubit assay on the PerkinElmer Victor X2. Illumina TruSeq Stranded total RNA-Seq Kit combined with 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 2nd strand synthesis, the cDNA:RNA hybrid is converted into to double-stranded cDNA (dscDNA) and dUTP incorporated into the 2nd 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 was first assessed on a Fragment Analyzer using the High Sensitivity DNA Kit (Agilent Technologies). All libraries were then pooled in equimolar ratio and sequenced. Sequencing was done at 50bp single-end reads. Once generated, sequence data was demultiplexed and Fastq files generated using Bcl2Fastq conversion software from Illumina.

SMRT-seq mapping and m^6A analysis

m^6A 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 50x coverage, sites with a phred-based quality score greater than 40 were marked as “1”, for strong evidence of methylation; sites with >=50x coverage but below QV40 were marked as “0” for unlikely to be methylated. For sites below 50x 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).

RNA-seq analysis and integration with methylomics

RNA-seq data was processed using the TrimGalore toolkit (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) which employs Cutadapt (62) to trim low quality bases and Illumina sequencing adapters from the 3’ end of the reads. Only reads that were 20nt 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 (63) using the STAR RNA-seq alignment tool (64). 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 was carried out using the DESeq2 (65) Bioconductor (66) package with the R statistical programming environment (https://www.R-project.org/).The false discovery rate was calculated to control for multiple hypothesis testing.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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 (a) the entire list of genes present in both analyses and (b) 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 number of differentially methylated genes) and observed differentially methylated and differentially expressed genes were compared. Fisher's 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 (67,68). 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 (69). The 14028s YhdJ protein sequence was used a 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/) (70,71). The PANTHER Overrepresentation Test was run using the *Salmonella* Typhimurium GO biological process reference, the test used Fisher's Exact test, and the correction was based on a calculated false discovery rate. All calculations were run automatically though 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 5mL of either LB or SPI-2 inducing media, and grown at 37˚C at 250RPM. OD600 measurements were taken every 30 minutes using a spectrophotometer (Pharmacia Biotech Novaspec II).

Gentamicin protection assay

Invasion and replication were measured as previously described (72-74). Briefly, bacteria were grown overnight, subcultured 1:33 into 1mL of LB, and grown for 2 hours and 45 minutes or until all strains entered late-log phase growth at 37˚C with 250 RPM. For any experiment using Δ*dam* bacteria, all bacteria were grown an extra 30 minutes (3 hours and 15 minutes) so that the Δ*dam* and Δ*dam*Δ*metJ* mutants reached late-exponential phase growth. 100,000 THP-1 monocytes, in antibiotic free media, were then infected by *S.* Typhimurium (MOI 5). At one hour post infection, cells were treated with gentamicin (50μg/mL), and IPTG was added 2 hours post infection to induce bacterial GFP expression. At 3 hours and 15 minutes post infection, cells were

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

read by a Guava Easycyte Plus flow cytometer (Millipore). At 22 hours and 45 minutes post infection, IPTG was added to remaining wells to induce GFP, and at 24 hours post infection, cells were quantified by flow cytometry. Percent host cell invasion was determined by quantifying the number of GFP+ cells 3 hours and 15 minutes post infection, and replication was assessed by determining the ratio of the median intensity of GFP positive cells at 24 hours post infection divided by the median of the GFP positive cells at 3 hours and 15 minutes post infection.

Motility assays

All strains were cultured overnight in LB, subcultured 1:33 into LB, and grown for 2 hours and 45 minutes at 37˚C with 250 RPM. A pipette tip was used to puncture and deliver 2μL of *S.* Typhimurium into the center of a 0.3% LB agar plate. Plates were incubated at 37°C for 6 hours 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 (7-14 weeks old) mice. Bacteria were grown overnight, subcultured 1:33, and grown for 2 hours and 45 minutes at 37˚C with 250 RPM. The bacteria were then washed and resuspended in PBS. Inoculums were confirmed by plating for CFUs. For oral infections, mice were fasted for 12 hours before infection, and given 100μL of a 10% sodium bicarbonate solution by oral gavage 30 minutes before infection. Mice then received a 1:1 mixture of two *S.* Typhimurium strains containing either pWSK29 (AmpR) or pWSK129 (KanR) (61), totaling 10^8 CFU in 100μL, by oral gavage. For intraperitoneal (IP) infections, mice were injected with a 1:1 mixture of two *S.* Typhimurium strains, totaling 10^3 CFU in 100uL, into the intraperitoneal space. For both models, tissues were harvested four days post infection, homogenized, and plated on LB agar containing either ampicillin or kanamycin. Competitive index was calculated as (# Strain A CFUs in tissue/# Strain B CFUs in tissue)/ (# Strain A CFUs in inoculum/# Strain B CFUs 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). 10μL reactions contained 5μL of the supermix, a final concentration of 500nM of each primer, and 2μL of cDNA. Reactions were run on a QuantStudio 3 thermo cycler. The cycling conditions were as follows: 95°C for 30 seconds, 40 cycles of 95 degrees for 15 seconds and 60°C for 60 seconds, and 60°C for 60 seconds. A melt curve was performed 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. Fold change represents $2^{-\Delta\Delta CT}$. Oligonucleotides are listed in **Supplemental Table 3**.

Western blotting

flhC was tagged with the 3xFLAG tag using recombineering as previously described (75). *S.* Typhimurium were grown overnight in LB, subcultured 1:33 in LB at 37˚C with 250 RPM until late log phase, and pelleted by centrifugation at 6,000xg for 5 minutes. Pellets were resuspended

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

in 2x laemmli buffer (Bio-Rad) with 5% 2-Mercaptoethanol, boiled for 10 minutes, 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-minute UV exposure. Following transfer onto Immun-Blot low-fluorescence PVDF membrane (Bio-Rad) using a Hoefer TE77X, blots were probed using an anti-FLAG M2 antibody (Sigma F3165). A florescent 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 seconds of UV exposure, and quantified using Fiji (76). The graphed relative signal is: (FLAG band intensity/Total Protein) divided by (FLAG band intensity in wild-type *flhC:FLAG3x* bacteria/Total Protein in wild-type *flhC:FLAG3x* bacteria).

Statistical analyses

Statistics were performed in Graphpad Prism 9 or Microsoft Excel, except where otherwise noted. Where noted inter-experimental 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 (average of all experiments divided by average of specific experiment). All statistical tests corresponding to reported p-values are described in the appropriate figure legends.

Results

A genome-wide screen to understand how growth conditions and methionine metabolism impact m^{6}A DNA methylation

While previous work on prokaryotic 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^{6}A DNA methylome changes under four biologically relevant conditions **(Figure 1A)**. We examined aerobic growth in LB media to late exponential phase, 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 (77). The second condition cultured bacteria in a minimal media used to induce expression of genes in the *Salmonella* Pathogenicity Island-2 (SPI-2) (58)—which include a type III secretion system turned on in the host cell to promote *Salmonella* vacuolar survival (78,79). The third and fourth conditions repeated growth in these media but used a methionine metabolism mutant *S.* Typhimurium strain, Δ*metJ* (80). The MetJ protein represses expression of methionine metabolism genes. Thus, Δ*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-adenosyl-methionine (SAM) (81), and the methyltransferase-inhibiting metabolites methylthioadenosine (82-84) and S-adenosyl-homocysteine (85-89). Of note, we have demonstrated that the Δ*metJ* mutant has attenuated SPI-1 secretion, motility, and virulence (90) and had previously hypothesized that these effects could be mediated through aberrant methylation.

In order to analyze the DNA methylome we performed PacBio SMRT-sequencing. This experiment, hereon called the "Methylation Experiment 1," was performed in 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 Δ*dam* and Δ*dam*Δ*metJ* mutants, which lack G(m^{6}A)TC (henceforth the * symbol will denote the adenosine that is m^{6}A

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

modified; 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. We also performed RNA-sequencing in biological triplicate on bacteria grown under these same conditions so that changes in DNA methylation could be correlated with changes in transcript abundance.

In total this experiment defined the methylation status of 61,704 adenosine bases **(GEO: GSE185578)**, however, methylation status of some bases under certain conditions could not be determined as coverage was below 50X coverage. 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. Additional methyl-bases were detected on the pWSK29 plasmid harbored in these strains, however, we did not include these bases in our analyses as this plasmid is not involved in the natural lifestyle of *S.* Typhimurium.

To compare methylation across conditions, we called methylation in two ways. First, we assigned each base a "percent methylated" value, which considered the percent of reads for each base that were counted as methylated compared to the total number of reads **(Supplemental File 1)**. However, as discussed later, these numbers are difficult to interpret without biological replicates and so we did not consider this quantitative measure for our initial analyses. Instead, we focused on a binary analysis which considered bases either methylated (if any methylation was predicted) or unmethylated **(Supplemental File 2)**. Using our binary analysis, we observed that there were nearly equal rates of m^6A methylation across wild-type and Δ*metJ* bacteria grown in LB and SPI-2 conditions (WT LB: 39,240 bases; WT SPI-2: 38,827 bases; Δ*metJ* SPI-1: 39,352 bases; Δ*metJ* SPI-2: 40,145 bases) **(Figure 1B)**. In contrast, Δ*dam* and Δ*dam*Δ*metJ* bacteria had significantly reduced methylation in LB (8,606 and 8,205 bases) and SPI-2-inducing conditions (9,924 and 9,378 bases) **(Supplemental Figure 1A)**.

There are subtle changes in methylation in response to changing conditions

We next examined how these bases were distributed across different methylation motifs. This analysis was primarily based on motifs that had been previously established for *S.* Typhimurium (91), though we were able to detect an additional motif, CRTA*YN_6CTC. Notably, two motifs (CAGA*G and GA*GN_6RTAYG) cannot always be distinguished, so we included bases that matched to both motifs in counts for each. Bases that did not map to any known motif were listed as "Other." As with the total amount of m^6A methylation, we found that in our four main conditions there were very few differences in the total numbers of most motifs **(Figure 1C)**, while deletion of *dam* had the predictable effect of nearly completely ablating methylation at the GA*TC motif **(Supplemental Figure 1B)**. A few bases in GA*TC motifs remained "methylated" following *dam* deletion; however, we hypothesize these bases represent miscalled methyl-bases.

The most notable change in motif abundance occurred at the ATGCA*T motif, which is methylated by the YhdJ methylase (92). We observed more ATGCA*T methylation in bacteria grown under SPI-2-inducing conditions ($p<0.00001$, Chi-Square Test) or in Δ*metJ* bacteria ($p<0.00001$, Chi-Square Test), with the highest ATGCA*T methylation present in Δ*metJ* bacteria grown under SPI-2-inducing conditions. In contrast, and consistent with previous reports in *E. coli* (32), we observed very little change in the total amount of GA*TC methylation present across these conditions. We also observed variation in the number of bases that mapped to the "Other" category, though it is unclear what this signifies.

After broadly characterizing the DNA methylome for each experimental condition, we next compared methylation at each individual base to identify differentially methylated bases (bases

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

that were called methylated in one condition but not another). The comparisons we used were wild-type *S.* Typhimurium grown in LB vs SPI-2-inducing conditions, wild-type vs Δ*metJ S.* Typhimurium grown in LB, and wild-type vs Δ*metJ S.* Typhimurium grown in SPI-2-inducing conditions. Interestingly, 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 **(Figure 2A, 2B, 2C)**. This demonstrates that while the methylome is slightly responsive to the environment and methionine metabolism, it remains largely static across strikingly different conditions.

In order to better understand the methylase(s) that drive these subtle changes across conditions, we next examined whether differentially methylated bases are enriched for any specific sequence motifs. We determined enrichment by comparing the frequency of each of the six motifs tested above in the differentially methylated sites against the frequency observed in the entire condition. For instance, for the wild-type LB vs SPI-2-inducing comparison, we examined the frequency of ATGCA*T methylation in the 423 bases uniquely methylated under SPI-2-inducing conditions and compared to the frequency of ATGCA*T in the 38,827 bases methylated under SPI-2-inducing conditions **(Figure 2D)**. 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 Δ*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-fold and 11-fold enrichment, accordingly) **(Figure 2E, 2F)**. Surprisingly, all other motifs were either present at similar or dramatically lower abundance among differentially methylated sites than expected by chance. This was particularly true for GA*TC methylation, which was depleted in differentially methylated bases in all conditions tested. This was perhaps expected given that Figure 1C demonstrated that overall GA*TC abundance did not change. However, we had hypothesized that global GA*TC patterning would differ across our conditions, as was observed with the *opvAB* operon in response to phage infection (48). Finally, we also saw significant enrichment of "other" motifs among differentially methylated bases (20 to 100-fold enrichment, depending on condition). This could suggest that undescribed methylases contribute to the flexible DNA methylome, or that "differentially methylated bases" are enriched for miscalled methyl-bases. Regardless, these data suggest that dynamic methylation, and in particular dynamic Dam methylation, is likely a relatively rare phenomenon.

A replication experiment demonstrates that SMRT-seq is highly reproducible and confirms differential methylation at YhdJ sites

While our first experiment provided evidence that there were modest differences in methylation across conditions, the reproducibility of these differences remained unclear. Many bacterial methylomics papers report methylation based on a single replicate, but we are unaware of any study that has demonstrated that this is sufficient. Because we were especially interested in whether methylation could explain any of the virulence related phenotypes that we observe with the Δ*metJ* mutant, specifically defects in SPI-1 secretion and motility, we chose to repeat our SMRT-seq experiment with wild-type and Δ*metJ* bacteria grown in LB **(Figure 3A;** Replication Methylation Experiment**)**. Further, to confirm that the significant enrichment in ATGCA*T methylation in Δ*metJ* bacteria we observed above **(Figure 1C, 2D)** was due to YhdJ, we also sequenced Δ*yhdJ* and Δ*yhdJ*Δ*metJ S.* Typhimurium grown in LB. Of note, while we sequenced eight samples across two SMRT cells in Methylation Experiment 1, here we sequenced these four samples on two SMRT cells, significantly increasing our sequencing depth. The resulting dataset

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

called the methylation status of 60,502 bases in at least one condition, and, strikingly, 60,501 of these bases were confidently called in all four conditions **(GEO: GSE185501**).

By assigning each base a binary value of methylated or unmethylated, we integrated our two datasets and examined the frequency at which bases that appeared in both datasets had the same methylation status in wild-type and Δ*metJ* bacteria. Reassuringly, using this binary measurement we found that ~97.5% of bases replicated their methylation status, demonstrating that our results were highly reproducible **(Figure 3B, Supplemental File 2)**. Importantly, we again observed that Δ*metJ* bacteria have increased ATGCA*T methylation and confirmed that YhdJ is the only ATGCA*T methylase active in both bacterial strains **(Figure 3C)**.

Next, we examined how reproducible changes in the percent methylation values were in our two experiments **(Supplemental File 1)**. This is an important measurement as previous work has speculated that methylation impacts bistable gene expression (48,93), and thus changes in the percent of the population in which a given base is methylated could have implications on the percent of the population expressing a given gene. While we saw considerable replication for this value in both wild-type (R^2=0.87) and Δ*metJ* (R^2=0.86), the noise in these data suggests that biological replicates are necessary to perform these analyses **(Figure 3D)**. Considering these experiments as separate biological replicates and using an arbitrary cutoff of 10% average differential methylation, we identified 2,528 sites (out of 50,962 total sites; 4.96%) that were differentially methylated between wild-type and Δ*metJ* bacteria using this quantitative method (**Figure 3E)**. 881 of these sites were more methylated in wild-type bacteria, and 1,647 were more methylated in Δ*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 dataset containing bases which were (a) reliably detected in wild-type and Δ*metJ* bacteria grown in LB in both experiments, and (b) were identically called methylated or unmethylated in both experiments. Using this dataset (52,594 bases), we determined which differentially methylated bases repeated across the two studies. Strikingly, while our data demonstrated most bases were called identically (**Figure 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. In fact, while there were 1,382 bases called differentially methylated in the first experiment **(Figure 2C)**, and 2,544 bases called differentially methylated in the replicate study **(Figure 3F)**, only 308 differentially methylated bases were identified in the combined dataset (**Figure 3G)**. Importantly, the overlap between these two replicates is significantly greater than expected by chance (3.7-fold enrichment; $p<0.0001$, one-tail binomial test), giving us high confidence in these 308 sites. These data once again emphasize the need for replicates in SMRT-seq based methylation experiments, even when performing binary analyses. Notably, while the absolute number of differentially methylated sites varied significantly across the three datasets (Experiment 1, Replication, and Combined), analysis of motifs among differentially methylated sites in the combined dataset revealed the same trends of enrichment in the ATGCA*T motif, but depletion for the GA*TC motif in unique Δ*metJ* sites **(Figure 3H)**.

Curiously, the enrichment we observed in "other" bases was significantly stronger in the combined dataset than we observed in either individual dataset, suggesting that these "other" bases truly represent a significant component of the dynamic DNA methylome. Using the Multiple Em for Motif Elicitation (MEME) software (94), we examined the 40 bases surrounding the 143 instances of "other" differential methylation in the Combined Dataset to identify shared motifs. This identified a single significant motif (E-value=6.1×10^{-16}), ACCWGG **(Supplemental Figure**

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

2A). The same motif was identified among the 969 differentially methylated "other" sites between LB and SPI-2 grown bacteria (Methylation Experiment 1; E-value = $5.5x10^{-206}$; **Supplemental Figure 2B**). Curiously, 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^5C motif CCWGG—methylated by Dcm. Together, this leads us to hypothesize that this dynamic "other" category is predominantly driven by changes in the flexible m^5C methylome, which may warrant further investigation using sequencing technologies better equipped to detect cytosine methylation.

***metJ* ablation leads to modest changes to the *S.* Typhimurium transcriptome**

Canonically, changes in DNA methylation are believed to lead to changes in transcription by enabling differential binding of transcription factors to genomic elements (reviewed (41)). However, studies that describe this in bacteria typically either (a) focus on single loci informed by methylase knockout experiments (for example, (48)), or (b) speculate on transcriptional control by methylation based on correlation of differential expression with static presence of methylation (for example (35)). No study has directly examined whether differential methylation across the *S.* Typhimurium genome correlates with differential expression. We attempted to fill this gap in knowledge by performing RNA-seq on bacteria grown under the same conditions used in our SMRT-seq experiments **(GEO: GSE185072, Supplemental File 3)**.

As expected, we identified many differentially expressed genes between wild-type bacteria grown in LB and SPI-2-inducing conditions **(Figure 4A)**, and a moderate number of genes differentially expressed between wild-type and Δ*metJ* bacteria **(Figure 4B, 4C)**. Notably, in striking contrast to the very modest differences in methylation that we observed between the LB and SPI-2 media grown bacteria (~3.2% of bases are differentially methylated across conditions, **Figure 2A**), 55.8% of the transcriptome changes across these two conditions. In order to facilitate this level of reprogramming, the cell must undergo dramatic changes in transcription factor binding and genomic organization. That the global m^6A methylome is resistant to this dynamic landscape represents impressive stability.

Further analysis of our Δ*metJ* differentially expressed genes revealed a number of expected and unexpected trends. For instance, during growth in LB, Δ*metJ* not only has the expected increase in methionine metabolism resulting from direct derepression of metabolism genes (80), but also decreased expression of motility genes as we had previously reported (90). While most of the reduced motility gene expression was expected, we were surprised to see a small reduction in *flhD* expression in Δ*metJ*, as we had not observed this in our previous work (90). However, we confirmed this result by qPCR and western blotting **(Supplemental Figure 3A, 3B)**, and speculate that improved DNase treatment in this study likely explains this difference. Examining the RNA-seq data further, we note that Δ*metJ S.* Typhimurium grown in both LB and SPI-2-inducing media have elevated expression of the oxidative stress response gene *katG* (95), as well as the putative iron-importer *sfbA* (96). This may suggest that the Δ*metJ* mutant is consistently under oxidative stress and/or iron starved. We also noted reduced expression of *rcsF* (a component of the Rcs two-component signaling system) and *ompF* (an outer membrane porin). Further, the most repressed gene in *metJ* grown in LB was *yfdH*, a component of SPI-16 which impacts form variation of the O12 antigen (97). Interestingly, under SPI-2 inducing conditions, we observed limited changes to the Δ*metJ* transcriptome beyond genes that were also differentially expressed in LB. Together, these changes in gene expression help generate hypotheses on why we observe virulence defects in Δ*metJ* bacteria (90).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Differential methylation does not correlate with differential expression

We next sought to integrate our differential expression data with our methylomics data **(Figure 4D)**. To do this, we considered genes that either (a) contained or (b) were the closest gene to one or more differentially methylated sites to be "differentially methylated genes." For each comparison, the status "differentially methylated" applied to the condition in which the methyl mark was present (*e.g.* when comparing LB grown vs SPI-2 grown bacteria, an LB-grown differentially methylated gene contains a methyl mark 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 differentially expressed genes among our differentially methylated genes, suggesting the two phenomena are not correlated across any of our comparisons (**Figure 4E, 4F, 4G)**.

We next examined whether stratification of our data could reveal enrichment of differentially expressed genes among a subclass of differentially methylated genes. While there are too few differentially expressed, differentially methylated genes in the Δ*metJ* comparisons to stratify the list further, we were able to perform these analyses on the wild-type LB vs SPI-2-media dataset. Specifically, we attempted to stratify our data by the differential expression direction of effect **(Supplemental Figure 4A, 4B)**, looking at only differentially methylated bases upstream of genes **(Supplemental Figure 4C)**, or the specifically examining differential YhdJ- or Dam-mediated differential methylation **(Supplemental Figure 4D, 4E)**. Regardless of our stratification, we did not observe any correlation between differential methylation and differential expression.

YhdJ is a dynamic, conserved methylase

While our data do not support a correlation between differential methylation and differential expression under these conditions, we hypothesized that within our datasets there may be discrete "exceptions" where differences in methylation *do* regulate transcriptional changes. To try to identify sites where methylation impacts expression, we first focused on the YhdJ methylase, which appears to be the most dynamic m^6A methylase in *S.* Typhimurium **(Figure 2D, 2E, 2F)**. YhdJ has been relatively understudied, as it is only experimentally described in one report as a non-essential methylase in *S.* Typhimurium and *E. coli* (92). To better understand YhdJ, we first examined its conservation across *Salmonella* serovars by aligning the sequences of over 9,000 strains from 10 serovars pulled from EnteroBase (67,68) against the *S.* Typhimurium 14028s protein sequence. While we observed some evidence of gene degradation in some isolated strains, only *Salmonella* Paratyphi A appeared to diverge from 14028s amino acid sequence significantly and consistently (most *S.* Paratyphi A strains harbor a 151* truncation), with most other strains having complete amino acid conservation **(Figure 5A)**. We hypothesized that YhdJ's conservation and its dynamic methylase activity are the result of it being an important regulator of *S.* Typhimurium physiology.

YhdJ plays little role in *Salmonella* physiology under standard laboratory conditions

To address whether YhdJ is a critical determinant of *S.* Typhimurium biology, we first tried to understand the nature of why YhdJ is more active in Δ*metJ* and SPI-2-media grown bacteria. To test whether this increased activity is due to increased *yhdJ* expression, we examined the *yhdJ* reads from our RNA-seq dataset above. Curiously, despite enhanced ATGCA*T methylation, wild-type *S.* Typhimurium grown in SPI-2 inducing conditions had reduced *yhdJ* expression compared to LB-grown bacteria **(Figure 5B)**. There were no differences in wild-type and Δ*metJ*

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

yhdJ expression. Thus, differences in YhdJ-mediated methylation are not due to differences in expression but may be due to differential access of the methylase to the genome.

To better understand YhdJ's cellular role, we examined which genes are targeted by YhdJ for methylation. To do this we examined four categories from our Methylation Experiment 1 dataset: (1) genes with ATGCA*T methylation under all conditions, (2) genes with ATGCA*T sites methylated in WT bacteria grown in SPI-2 inducing conditions, but not LB, (3) genes with ATGCA*T sites methylated in Δ*metJ* LB-grown bacteria, but not wild-type, and (4) genes with ATGCA*T sites methylated in Δ*metJ* SPI-2-induced bacteria, but not wild-type. We utilized GO analysis (70,71) and found that all four gene lists were enriched for a plethora of metabolic genes (**Supplemental Table 4)**. Curiously, the list of GO-terms did not fundamentally change across our various lists. Despite these changes to metabolic genes, we confirmed previous reports that YhdJ is not required for growth in LB **(Figure 5C)** (92), nor growth in SPI-2-inducing media **(Figure 5D).**

We next examined whether *yhdJ* had any functional impacts on *Salmonella* performance during assays relevant to LB- or SPI-2-grown bacteria. First, using a modified flow-cytometry based gentamicin protection assay (72,73), we examined whether YhdJ contributes to *S.* Typhimurium entry into or replication within THP-1 monocytes. Surprisingly, we found that *yhdJ* deletion subtly and statistically significantly increased the percent of host cells infected compared to wild-type bacteria **(Figure 5E)**. We also tested whether there were any interactions with the previously observed Δ*metJ* invasion defect (90), but found no significant interaction. Next, we examined bacterial replication using our flow cytometric approach (74), but found no effects of *yhdJ* or *metJ* on replication **(Figure 5F)**. Finally, we tested whether YhdJ was required for bacterial motility or for the effects of *metJ* deletion on motility but found no effect **(Figure 5G).**

Despite the absence of effects in these *in vitro* systems, we next tested YhdJ's role in murine virulence. First, we infected C57BL/6J mice by intraperitoneal injection with equal amounts of wild-type and Δ*yhdJ* bacteria and found that we were able to recover near-equal amounts of bacteria from spleens four days post infection **(Table 1)**, demonstrating that YhdJ does not contribute to *S.* Typhimurium fitness in this model. We also tested whether Δ*yhdJ* had a genetic interaction with Δ*metJ* but found no difference in competitive indexes between Δ*yhdJ* and Δ*yhdJ*Δ*metJ* (Competitive Index = 0.26) and between wild-type and Δ*metJ* (Competitive Index = 0.23) **(Table 1)**. Next, we used the enteric fever model of infection where mice are infected orally without antibiotic pretreatment and found that Δ*yhdJ S.* Typhimurium slightly outcompeted their wild-type counterparts in the spleen (Competitive Index = 1.65), though statistically significant differences were not observed in the ileum **(Table 1)**. Together, our *in vitro* and *in vivo* data suggest that despite evolutionary conservation, *yhdJ* is either not involved in or slightly detrimental to *Salmonella* virulence processes.

Our inability to detect phenotypes associated with Δ*yhdJ* led us to question whether YhdJ methylation has any impact on the transcriptome under these conditions. To address this, we performed RNA-seq on wild-type and Δ*yhdJ* bacteria grown in LB or SPI-2 inducing conditions **(GEO: GSE185073, Supplemental File 4)**. Strikingly, knocking out *yhdJ* had almost no impact on the transcriptome, despite complete loss of methylation at the ATGCA*T motif under all conditions tested **(Figure 3C)**. Apart from *yhdJ* itself, only 12 genes were differentially expressed between wild-type and Δ*yhdJ* bacteria grown in LB (**Figure 5I, Table 2)**, and no genes were differentially expressed under SPI-2 inducing conditions **(Figure 5J)**. Curiously, the differentially expressed genes are enriched for de novo UMP biosynthetic processes (False Discovery Rate=6.3×10^{-4}) and de novo pyrimidine nucleobase biosynthetic processes (False Discovery

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Rate=8.31x10^{-4}) (70,71). 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 Dataset **(Table 2)**. Even this methyl-base is present near the 3' end of the gene and so it is not immediately obvious if or how methylation at this base could impact gene expression. Based on these data we hypothesize that either (a) YhdJ has modest and indirect impacts on gene expression, or (b) off-target impacts of recombineering led to altered gene expression in this mutant. Regardless, our data once again argue against the hypothesis that an environmentally sensitive methylome has widespread and direct impacts on gene expression, as our most dynamic methylation motif appears to have no direct impact on transcript abundance.

Despite limited changes to the GA*TC methylome, *metJ* and *dam* interact to influence *S.* Typhimurium invasion and motility

While we observed significant depletion of the GA*TC motif among our differentially methylated bases **(Figure 2)**, the rich history of Dam methylation impacting *Salmonella* biology led us to hypothesize that the small number of differentially GA*TC methylated genes could have important transcriptional consequences. To test this hypothesis, we examined whether *dam* is required for the impacts of *metJ* deletion on host cell invasion and motility. We first tested whether there was an interaction between Δ*dam* and Δ*metJ* on THP-1 invasion. Importantly, we observe a small growth defect in Δ*dam* and Δ*dam*Δ*metJ* bacteria, and so all bacteria used for these experiments were grown an extra 30 minutes prior to infection to standardize the growth phase used. Knocking out *dam* only modestly reduced the impact of Δ*metJ* on invasion and is therefore unlikely to be the primary mechanism by which *metJ* deletion impacts invasion **(Figure 6A)**. In contrast, the impairment in motility caused by *metJ* deletion was completely abrogated in Δ*dam* genetic background, suggesting that *dam* and *metJ* impact motility through the same pathway **(Figure 6B)**. Importantly, complementation of *dam* on a low copy number plasmid (pWSK129) restored differences between Δ*dam* and Δ*dam*Δ*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).

We hypothesized that the genetic interaction between Dam and MetJ could signify that differential GA*TC methylation in the Δ*metJ* mutant suppresses bacterial motility. In striking contrast to this hypothesis, our combined dataset revealed no genes that were both differentially GA*TC methylated and expressed (except for the deleted *metJ* itself) **(Table 3)**. 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 datasets at all GA*TC methylated sites in which the nearest gene is differentially expressed identified 17 sites that had a ≥10% average difference in methylation between wild-type and Δ*metJ* bacteria **(Figure 6C).** Restricting our search to GA*TC sites that are upstream of differentially expressed genes further refined this list to 2 sites, specifically, both strands of a single GATC motif upstream of the chemotaxis gene *tsr* that shows elevated methylation in Δ*metJ* **(Figure 6D and 6E, Table 4).**

This hypermethylation led us to hypothesize that increased methylation upstream of *tsr* in Δ*metJ* 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 Δ*metJ* to drive a motility defect **(Figure 6F,** interaction term p=0.005**)**. Additionally, a search for the methylation-sensitive transcription factor CRP (41) binding motif (AAATGTGATCTAGATCACATTT) in the

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

tsr promoter with the MEME FIMO Tool (98) demonstrated that the hypermethylated residue lies within a putative CRP binding site. Together, these data tentatively support a model in which increased methylation upstream of *tsr* in Δ*metJ* 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 Δ*metJ* motility defect

Given that unlike Δ*dam* (**Figure 6B),** Δ*tsr* does not account for the entire impact of *metJ* deletion on motility **(Figure 6F**, Ratio of Δ*tsr::kan*Δ*metJ*/Δ*tsr::kan* = 0.9**)**, we hypothesized that there may be additional differences in GA*TC methylation between wild-type and Δ*metJ* bacteria that impact motility. Further examination of our quantitative methylation dataset **(Table 4)** revealed one additional plausible hypothesis: a site in the *flhDC* promoter (-278) that barely missed our 10% threshold (9.5% more methylated in Δ*metJ* bacteria). We decided to test this site as well, as 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 Δ*metJ*, we performed site directed mutagenesis on the *S.* Typhimurium chromosome to mutate the base from GATC to GTTC. However, this mutation had no effect on motility in wild-type or Δ*metJ* bacteria **(Supplemental Figure 5)**, disproving the hypothesis that this site could contribute to the Δ*dam* epistatic effect.

Discussion

Together, our data paint a narrative in which the *S.* Typhimurium environment, epigenome, and transcriptome are, with exception, decoupled. While previous work has identified examples of DNA methylation influencing gene expression, this phenomenon seems to be relatively rare for this microbe. In contrast to previous work on m^5C methylation in *V. cholera* (11), we did not observe enrichment of differentially expressed genes among differentially methylated genes across our conditions. Further, 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 non-essential 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). While 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 3D conformation of the genome that prevent interactions between YhdJ and its motif.

We propose three potential explanations for the absence of a correlation between global m^6A DNA methylation and gene expression in our data. The first is that while *S.* Typhimurium can and do use m^6A 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 (12,21,22,29,30,36,53-57,99), they are rare exceptions to the general rules of *S.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Typhimurium gene regulation. The second hypothesis is that the four conditions tested here (wild-type or Δ*metJ* bacteria grown in LB or SPI-2 inducing media) are non-representative conditions where methylation plays little impact on transcription. Notably, while this is possible, these conditions were specifically chosen as they are (a) relevant to the pathogenic capacity of the bacteria, (b) the conditions most frequently studied in laboratory settings, or (c) 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 observed *S.* Typhimurium phenomenon are not linked to changes in m^6A methylation. The third possibility is that while m^6A is the most common modification to the *S.* Typhimurium genome, other modifications (m^5C, phosphorothioation, *etc*.) may have more important impacts on gene expression.

In addition to our conceptual insights about methylation and gene expression, our experimental design enabled us to evaluate the reproducibility of SMRT-seq. Notably, many of the previous studies on m^6A methylation have been performed in biological singlet under the assumption each sequenced genome from each independent bacterium in the population is, in itself, a sort of replicate. In our work, while we found relatively high reproducibility in both the binary assignment of bases as methylated or unmethylated and the quantitative percent methylation of each base, some impactful noise was observed, especially in differentially methylated bases. Our conclusion from these experiments is that biological singlet is almost certainly insufficient when examining dynamic changes in methylomic experiments.

Despite finding that there are not correlations between global DNA methylation and gene expression, we were able to pair our datasets with data from Δ*dam* strains to identify a GA*TC locus upstream of *tsr* that is potentially partially responsible for the motility defect in Δ*metJ* bacteria. In contrast, we found that a second plausible hypothesis, that a GATC site upstream of the master motility regulator *flhDC* regulates expression, did not impact the Δ*metJ* phenotype. This latter finding should serve as a warning to the methylation field, as we note a disturbing and growing trend in which researchers attempt to link genetic data from static whole-genome methylation studies to phenotypic data from Δ*dam S.* Typhimurium. While the results from these experiments can be tantalizing, the pleotropic nature of Dam makes them difficult to interpret. Similarly, experiments that only leverage global transcription factor knockouts (*eg.* Δ*crp*) should also be considered inconclusive because of the definitionally far reaching consequences of those deletions. Instead, if researchers wish to comment on the impacts of specific GA*TC motifs on natural phenomena, direct manipulation of those motifs is required.

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. 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^6A methylation in *S.* Typhimurium.

Data Availability

All sequencing data is available in the NCBI's Gene Expression Omnibus (GEO) (100) Super Series (GSE185077). This includes both SMRT-seq experiments (GSE185578 and GSE185501), as well as both RNA-seq experiments (GSE185072 and GSE185073). All biological resources are available upon request to Dr. Dennis Ko.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Acknowledgements

The authors would like to 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 Dr. 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, and Rachel Keener for their helpful discussions about the manuscript. Finally, we are grateful to Dr. 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.

Funding

This work was supported by the National Institutes of Health [1F31AI143147 to JSB, R01AI118903 to DCK, R21AI144586 to DCK]. The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

No authors report a conflict of interest.

References

1. Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul, C.L. (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.
2. Loenen, W.A., Dryden, D.T., Raleigh, E.A., Wilson, G.G. and Murray, N.E. (2014) Highlights of the DNA cutters: a short history of the restriction enzymes. *Nucleic Acids Res*, **42**, 3-19.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

3. Glickman, B.W. and Radman, M. (1980) Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. *Proc Natl Acad Sci U S A*, **77**, 1063-1067.
4. Pukkila, P.J., Peterson, J., Herman, G., Modrich, P. and Meselson, M. (1983) Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in Escherichia coli. *Genetics*, **104**, 571-582.
5. Glickman, B.W. (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.
6. Schlagman, S.L., Hattman, S. and Marinus, M.G. (1986) Direct role of the Escherichia coli Dam DNA methyltransferase in methylation-directed mismatch repair. *Journal of bacteriology*, **165**, 896-900.
7. Torreblanca, J. and Casadesus, J. (1996) DNA adenine methylase mutants of Salmonella typhimurium and a novel dam-regulated locus. *Genetics*, **144**, 15-26.
8. Robinson, V.L., Oyston, P.C. and Titball, R.W. (2005) A dam mutant of Yersinia pestis is attenuated and induces protection against plague. *FEMS Microbiol Lett*, **252**, 251-256.
9. Watson, M.E., Jr., Jarisch, J. and Smith, A.L. (2004) Inactivation of deoxyadenosine methyltransferase (dam) attenuates Haemophilus influenzae virulence. *Mol Microbiol*, **53**, 651-664.
10. Julio, S.M., Heithoff, D.M., Provenzano, D., Klose, K.E., Sinsheimer, R.L., Low, D.A. and Mahan, M.J. (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.
11. Chao, M.C., Zhu, S., Kimura, S., Davis, B.M., Schadt, E.E., Fang, G. and Waldor, M.K. (2015) A Cytosine Methyltransferase Modulates the Cell Envelope Stress Response in the Cholera Pathogen [corrected]. *PLoS Genet*, **11**, e1005666.
12. Collier, J., McAdams, H.H. and Shapiro, L. (2007) A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proc Natl Acad Sci U S A*, **104**, 17111-17116.
13. Stephens, C., Reisenauer, A., Wright, R. and Shapiro, L. (1996) A cell cycle-regulated bacterial DNA methyltransferase is essential for viability. *Proc Natl Acad Sci U S A*, **93**, 1210-1214.
14. Zweiger, G., Marczynski, G. and Shapiro, L. (1994) A Caulobacter DNA methyltransferase that functions only in the predivisional cell. *J Mol Biol*, **235**, 472-485.
15. Campbell, J.L. and Kleckner, N. (1990) E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell*, **62**, 967-979.
16. Campellone, K.G., Roe, A.J., Lobner-Olesen, A., Murphy, K.C., Magoun, L., Brady, M.J., Donohue-Rolfe, A., Tzipori, S., Gally, D.L., Leong, J.M. *et al.* (2007) Increased adherence and actin pedestal formation by dam-deficient enterohaemorrhagic Escherichia coli O157:H7. *Mol Microbiol*, **63**, 1468-1481.
17. Erova, T.E., Pillai, L., Fadl, A.A., Sha, J., Wang, S., Galindo, C.L. and Chopra, A.K. (2006) DNA adenine methyltransferase influences the virulence of Aeromonas hydrophila. *Infect Immun*, **74**, 410-424.
18. Mehling, J.S., Lavender, H. and Clegg, S. (2007) A Dam methylation mutant of Klebsiella pneumoniae is partially attenuated. *FEMS Microbiol Lett*, **268**, 187-193.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

19. Taylor, V.L., Titball, R.W. and Oyston, P.C.F. (2005) Oral immunization with a dam mutant of Yersinia pseudotuberculosis protects against plague. *Microbiology (Reading)*, **151**, 1919-1926.
20. Henderson, I.R. and 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. *Journal of bacteriology*, **181**, 2132-2141.
21. Cota, I., Blanc-Potard, A.B. and 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.
22. Broadbent, S.E., Davies, M.R. and van der Woude, M.W. (2010) Phase variation controls expression of Salmonella lipopolysaccharide modification genes by a DNA methylation-dependent mechanism. *Mol Microbiol*, **77**, 337-353.
23. Falker, S., Schilling, J., Schmidt, M.A. and Heusipp, G. (2007) Overproduction of DNA adenine methyltransferase alters motility, invasion, and the lipopolysaccharide O-antigen composition of Yersinia enterocolitica. *Infect Immun*, **75**, 4990-4997.
24. Sarnacki, S.H., Marolda, C.L., Noto Llana, M., Giacomodonato, M.N., Valvano, M.A. and Cerquetti, M.C. (2009) Dam methylation controls O-antigen chain length in Salmonella enterica serovar enteritidis by regulating the expression of Wzz protein. *Journal of bacteriology*, **191**, 6694-6700.
25. Sarnacki, S.H., Castaneda Mdel, R., Noto Llana, M., Giacomodonato, M.N., Valvano, M.A. and Cerquetti, M.C. (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.
26. Murphy, K.C., Ritchie, J.M., Waldor, M.K., Lobner-Olesen, A. and Marinus, M.G. (2008) Dam methyltransferase is required for stable lysogeny of the Shiga toxin (Stx2)-encoding bacteriophage 933W of enterohemorrhagic Escherichia coli O157:H7. *Journal of bacteriology*, **190**, 438-441.
27. Gordeeva, J., Morozova, N., Sierro, N., Isaev, A., Sinkunas, T., Tsvetkova, K., Matlashov, M., Truncaite, L., Morgan, R.D., Ivanov, N.V. *et al.* (2019) BREX system of Escherichia coli distinguishes self from non-self by methylation of a specific DNA site. *Nucleic Acids Res*, **47**, 253-265.
28. Torreblanca, J., Marques, S. and Casadesus, J. (1999) Synthesis of FinP RNA by plasmids F and pSLT is regulated by DNA adenine methylation. *Genetics*, **152**, 31-45.
29. Camacho, E.M. and Casadesus, J. (2005) Regulation of traJ transcription in the Salmonella virulence plasmid by strand-specific DNA adenine hemimethylation. *Mol Microbiol*, **57**, 1700-1718.
30. Blyn, L.B., Braaten, B.A. and Low, D.A. (1990) Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. *EMBO J*, **9**, 4045-4054.
31. Balbontin, R., Rowley, G., Pucciarelli, M.G., Lopez-Garrido, J., Wormstone, Y., Lucchini, S., Garcia-Del Portillo, F., Hinton, J.C. and Casadesus, J. (2006) DNA adenine methylation regulates virulence gene expression in Salmonella enterica serovar Typhimurium. *Journal of bacteriology*, **188**, 8160-8168.
32. Cohen, N.R., Ross, C.A., Jain, S., Shapiro, R.S., Gutierrez, A., Belenky, P., Li, H. and Collins, J.J. (2016) A role for the bacterial GATC methylome in antibiotic stress survival. *Nat Genet*, **48**, 581-586.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

33. Shell, S.S., Prestwich, E.G., Baek, S.H., Shah, R.R., Sassetti, C.M., Dedon, P.C. and Fortune, S.M. (2013) DNA methylation impacts gene expression and ensures hypoxic survival of Mycobacterium tuberculosis. *PLoS Pathog*, **9**, e1003419.
34. Badie, G., Heithoff, D.M., Sinsheimer, R.L. and Mahan, M.J. (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. *Journal of bacteriology*, **189**, 1556-1564.
35. Oliveira, P.H., Ribis, J.W., Garrett, E.M., Trzilova, D., Kim, A., Sekulovic, O., Mead, E.A., Pak, T., Zhu, S., Deikus, G. *et al.* (2020) Epigenomic characterization of Clostridioides difficile finds a conserved DNA methyltransferase that mediates sporulation and pathogenesis. *Nat Microbiol*, **5**, 166-180.
36. Brunet, Y.R., Bernard, C.S., Gavioli, M., Lloubes, R. and 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.
37. Chen, L., Paulsen, D.B., Scruggs, D.W., Banes, M.M., Reeks, B.Y. and Lawrence, M.L. (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.
38. Heithoff, D.M., Sinsheimer, R.L., Low, D.A. and Mahan, M.J. (1999) An essential role for DNA adenine methylation in bacterial virulence. *Science*, **284**, 967-970.
39. Garcia-Del Portillo, F., Pucciarelli, M.G. and 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.
40. Atack, J.M., Tan, A., Bakaletz, L.O., Jennings, M.P. and Seib, K.L. (2018) Phasevarions of Bacterial Pathogens: Methylomics Sheds New Light on Old Enemies. *Trends Microbiol*, **26**, 715-726.
41. Sanchez-Romero, M.A. and Casadesus, J. (2020) The bacterial epigenome. *Nat Rev Microbiol*, **18**, 7-20.
42. Beaulaurier, J., Schadt, E.E. and Fang, G. (2019) Deciphering bacterial epigenomes using modern sequencing technologies. *Nat Rev Genet*, **20**, 157-172.
43. Marinus, M.G. and Lobner-Olesen, A. (2014) DNA Methylation. *EcoSal Plus*, **6**.
44. Flusberg, B.A., Webster, D.R., Lee, J.H., Travers, K.J., Olivares, E.C., Clark, T.A., Korlach, J. and Turner, S.W. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature methods*, **7**, 461-465.
45. Rand, A.C., Jain, M., Eizenga, J.M., Musselman-Brown, A., Olsen, H.E., Akeson, M. and Paten, B. (2017) Mapping DNA methylation with high-throughput nanopore sequencing. *Nature methods*, **14**, 411-413.
46. Tourancheau, A., Mead, E.A., Zhang, X.S. and Fang, G. (2021) Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nature methods*, **18**, 491-498.
47. Roberts, R.J., Vincze, T., Posfai, J. and Macelis, D. (2015) REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res*, **43**, D298-299.
48. Cota, I., Bunk, B., Sproer, C., Overmann, J., Konig, C. and Casadesus, J. (2016) OxyR-dependent formation of DNA methylation patterns in OpvABOFF and OpvABON cell lineages of Salmonella enterica. *Nucleic Acids Res*, **44**, 3595-3609.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

49. Erill, I., Puigvert, M., Legrand, L., Guarischi-Sousa, R., Vandecasteele, C., Setubal, J.C., Genin, S., Guidot, A. and Valls, M. (2017) Comparative Analysis of Ralstonia solanacearum Methylomes. *Front Plant Sci*, **8**, 504.
50. Modlin, S.J., Conkle-Gutierrez, D., Kim, C., Mitchell, S.N., Morrissey, C., Weinrick, B.C., Jacobs, W.R., Ramirez-Busby, S.M., Hoffner, S.E. and Valafar, F. (2020) Drivers and sites of diversity in the DNA adenine methylomes of 93 Mycobacterium tuberculosis complex clinical isolates. *Elife*, **9**.
51. Kahramanoglou, C., Prieto, A.I., Khedkar, S., Haase, B., Gupta, A., Benes, V., Fraser, G.M., Luscombe, N.M. and Seshasayee, A.S. (2012) Genomics of DNA cytosine methylation in Escherichia coli reveals its role in stationary phase transcription. *Nat Commun*, **3**, 886.
52. Davis-Richardson, A.G., Russell, J.T., Dias, R., McKinlay, A.J., Canepa, R., Fagen, J.R., Rusoff, K.T., Drew, J.C., Kolaczkowski, B., Emerich, D.W. *et al.* (2016) Integrating DNA Methylation and Gene Expression Data in the Development of the Soybean-Bradyrhizobium N2-Fixing Symbiosis. *Front Microbiol*, **7**, 518.
53. Peterson, S.N. and Reich, N.O. (2008) Competitive Lrp and Dam assembly at the pap regulatory region: implications for mechanisms of epigenetic regulation. *J Mol Biol*, **383**, 92-105.
54. Hernday, A., Krabbe, M., Braaten, B. and Low, D. (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci U S A*, **99 Suppl 4**, 16470-16476.
55. Waldron, D.E., Owen, P. and Dorman, C.J. (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.
56. Wallecha, A., Munster, V., Correnti, J., Chan, T. and van der Woude, M. (2002) Dam- and OxyR-dependent phase variation of agn43: essential elements and evidence for a new role of DNA methylation. *Journal of bacteriology*, **184**, 3338-3347.
57. Garcia-Pastor, L., Sanchez-Romero, M.A., Jakomin, M., Puerta-Fernandez, E. and 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.
58. Coombes, B.K., Brown, N.F., Valdez, Y., Brumell, J.H. and Finlay, B.B. (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.
59. Datsenko, K.A. and Wanner, B.L. (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.
60. Yang, J., Sun, B., Huang, H., Jiang, Y., Diao, L., Chen, B., Xu, C., Wang, X., Liu, J., Jiang, W. *et al.* (2014) High-efficiency scarless genetic modification in Escherichia coli by using lambda red recombination and I-SceI cleavage. *Appl Environ Microbiol*, **80**, 3826-3834.
61. Wang, R.F. and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli. *Gene*, **100**, 195-199.
62. Hwang, S., Martinez, D., Perez, P. and Rinaldi, C. (2011) Effect of surfactant-coated iron oxide nanoparticles on the effluent water quality from a simulated sequencing batch reactor treating domestic wastewater. *Environ Pollut*, **159**, 3411-3415.
63. Kersey, P.J., Staines, D.M., Lawson, D., Kulesha, E., Derwent, P., Humphrey, J.C., Hughes, D.S., Keenan, S., Kerhornou, A., Koscielny, G. *et al.* (2012) Ensembl Genomes:

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

an integrative resource for genome-scale data from non-vertebrate species. *Nucleic Acids Res*, **40**, D91-97.

64. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.
65. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 550.
66. Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T. *et al.* (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nature methods*, **12**, 115-121.
67. Alikhan, N.F., Zhou, Z., Sergeant, M.J. and Achtman, M. (2018) A genomic overview of the population structure of Salmonella. *PLoS Genet*, **14**, e1007261.
68. Achtman, M., Zhou, Z., Alikhan, N.F., Tyne, W., Parkhill, J., Cormican, M., Chiou, C.S., Torpdahl, M., Litrup, E., Prendergast, D.M. *et al.* (2020) Genomic diversity of Salmonella enterica -The UoWUCC 10K genomes project. *Wellcome Open Res*, **5**, 223.
69. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.
70. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, **25**, 25-29.
71. Gene Ontology, C. (2021) The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res*, **49**, D325-D334.
72. Alvarez, M.I., Glover, L.C., Luo, P., Wang, L., Theusch, E., Oehlers, S.H., Walton, E.M., Tram, T.T.B., Kuang, Y.L., Rotter, J.I. *et al.* (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.
73. Jaslow, S.L., Gibbs, K.D., Fricke, W.F., Wang, L., Pittman, K.J., Mammel, M.K., Thaden, J.T., Fowler, V.G., Jr., Hammer, G.E., Elfenbein, J.R. *et al.* (2018) Salmonella Activation of STAT3 Signaling by SarA Effector Promotes Intracellular Replication and Production of IL-10. *Cell Rep*, **23**, 3525-3536.
74. Wang, L., Pittman, K.J., Barker, J.R., Salinas, R.E., Stanaway, I.B., Williams, G.D., Carroll, R.J., Balmat, T., Ingham, A., Gopalakrishnan, A.M. *et al.* (2018) An Atlas of Genetic Variation Linking Pathogen-Induced Cellular Traits to Human Disease. *Cell Host Microbe*, **24**, 308-323 e306.
75. Uzzau, S., Figueroa-Bossi, N., Rubino, S. and Bossi, L. (2001) Epitope tagging of chromosomal genes in Salmonella. *Proc Natl Acad Sci U S A*, **98**, 15264-15269.
76. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nature methods*, **9**, 676-682.
77. Galan, J.E. and Curtiss, R., 3rd. (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.
78. Shea, J.E., Hensel, M., Gleeson, C. and Holden, D.W. (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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

79. Ochman, H., Soncini, F.C., Solomon, F. and Groisman, E.A. (1996) Identification of a pathogenicity island required for Salmonella survival in host cells. *Proc Natl Acad Sci U S A*, **93**, 7800-7804.
80. Hondorp, E.R. and Matthews, R.G. (2006) Methionine. *EcoSal Plus*, **2**.
81. Cantoni, G.L. (1953) S-Adenosylmethionine; a new intermediate formed enzymatically from L-methionine and adenosinetriphosphate. *J Biol Chem*, **204**, 403-416.
82. Bigaud, E. and Corrales, F.J. (2016) Methylthioadenosine (MTA) Regulates Liver Cells Proteome and Methylproteome: Implications in Liver Biology and Disease. *Mol Cell Proteomics*, **15**, 1498-1510.
83. Pascale, R.M., Simile, M.M., Satta, G., Seddaiu, M.A., Daino, L., Pinna, G., Vinci, M.A., Gaspa, L. and 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.
84. Woodcock, D.M., Adams, J.K., Allan, R.G. and Cooper, I.A. (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.
85. Yi, P., Melnyk, S., Pogribna, M., Pogribny, I.P., Hine, R.J. and James, S.J. (2000) Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. *J Biol Chem*, **275**, 29318-29323.
86. Hoffman, D.R., Cornatzer, W.E. and Duerre, J.A. (1979) Relationship between tissue levels of S-adenosylmethionine, S-adenylhomocysteine, and transmethylation reactions. *Can J Biochem*, **57**, 56-65.
87. Zappia, V., Zydek-Cwick, R. and Schlenk, F. (1969) The specificity of S-adenosylmethionine derivatives in methyl transfer reactions. *J Biol Chem*, **244**, 4499-4509.
88. Mull, L., Ebbs, M.L. and Bender, J. (2006) A histone methylation-dependent DNA methylation pathway is uniquely impaired by deficiency in Arabidopsis S-adenosylhomocysteine hydrolase. *Genetics*, **174**, 1161-1171.
89. Barber, J.R. and Clarke, S. (1984) Inhibition of protein carboxyl methylation by S-adenosyl-L-homocysteine in intact erythrocytes. Physiological consequences. *J Biol Chem*, **259**, 7115-7122.
90. Bourgeois, J.S., Zhou, D., Thurston, T.L.M., Gilchrist, J.J. and Ko, D.C. (2018) Methylthioadenosine Suppresses Salmonella Virulence. *Infect Immun*, **86**.
91. Pirone-Davies, C., Hoffmann, M., Roberts, R.J., Muruvanda, T., Timme, R.E., Strain, E., Luo, Y., Payne, J., Luong, K., Song, Y. *et al.* (2015) Genome-wide methylation patterns in Salmonella enterica Subsp. enterica Serovars. *PLoS One*, **10**, e0123639.
92. Broadbent, S.E., Balbontin, R., Casadesus, J., Marinus, M.G. and van der Woude, M. (2007) YhdJ, a nonessential CcrM-like DNA methyltransferase of Escherichia coli and Salmonella enterica. *Journal of bacteriology*, **189**, 4325-4327.
93. Sanchez-Romero, M.A., Olivenza, D.R., Gutierrez, G. and Casadesus, J. (2020) Contribution of DNA adenine methylation to gene expression heterogeneity in Salmonella enterica. *Nucleic Acids Res*.
94. Bailey, T.L., Johnson, J., Grant, C.E. and Noble, W.S. (2015) The MEME Suite. *Nucleic Acids Res*, **43**, W39-49.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

95. Loewen, P.C. and Stauffer, G.V. (1990) Nucleotide sequence of katG of Salmonella typhimurium LT2 and characterization of its product, hydroperoxidase I. *Mol Gen Genet*, **224**, 147-151.
96. Pattery, T., Hernalsteens, J.P. and De Greve, H. (1999) Identification and molecular characterization of a novel Salmonella enteritidis pathogenicity islet encoding an ABC transporter. *Mol Microbiol*, **33**, 791-805.
97. Bogomolnaya, L.M., Santiviago, C.A., Yang, H.J., Baumler, A.J. and Andrews-Polymenis, H.L. (2008) 'Form variation' of the O12 antigen is critical for persistence of Salmonella Typhimurium in the murine intestine. *Mol Microbiol*, **70**, 1105-1119.
98. Grant, C.E., Bailey, T.L. and Noble, W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017-1018.
99. Sanchez-Romero, M.A., Olivenza, D.R., Gutierrez, G. and Casadesus, J. (2020) Contribution of DNA adenine methylation to gene expression heterogeneity in Salmonella enterica. *Nucleic Acids Res*, **48**, 11857-11867.
100. Edgar, R., Domrachev, M. and Lash, A.E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*, **30**, 207-210.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figures and Tables

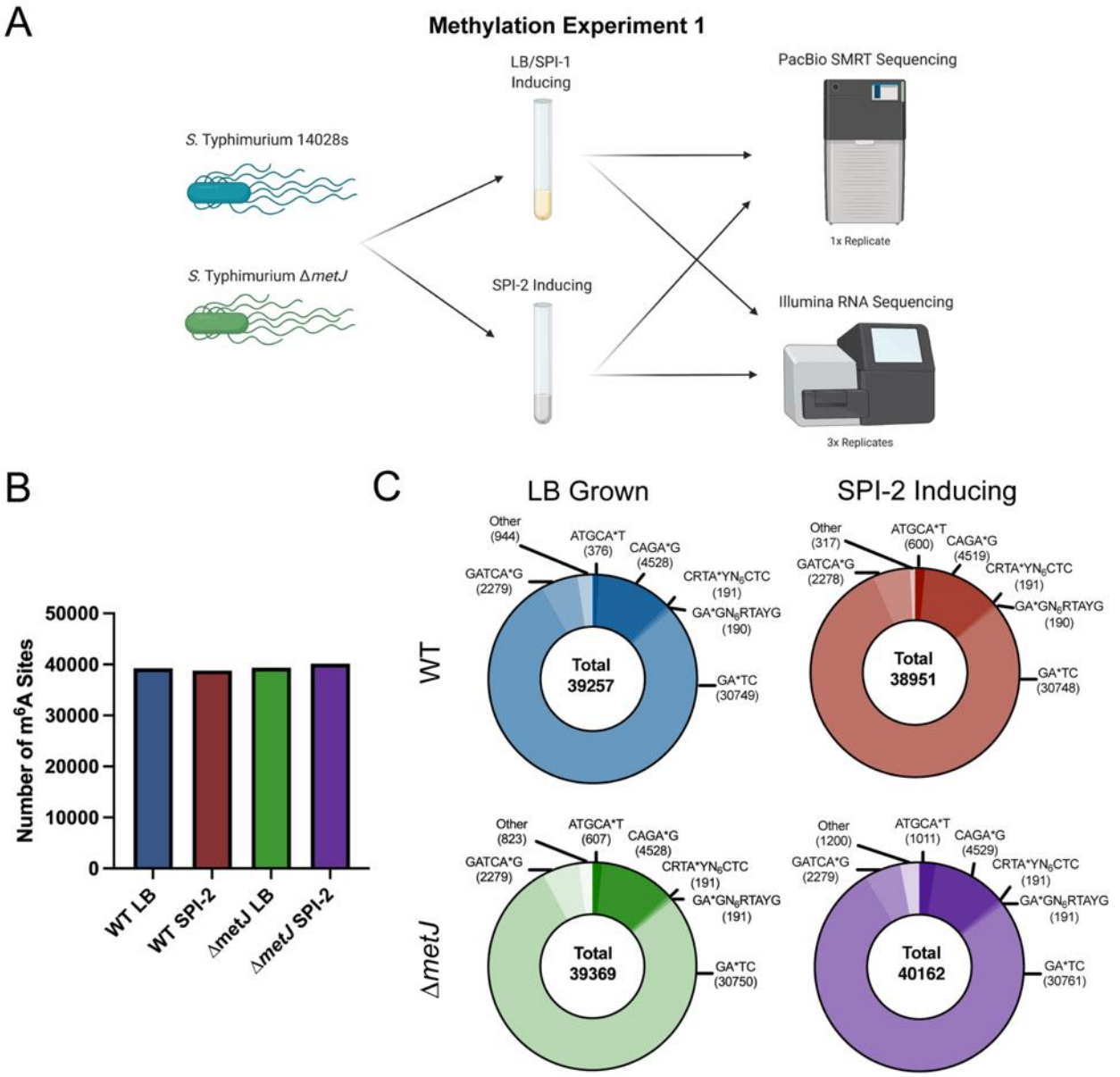


Figure 1: Genome-wide analysis of m^{6}A DNA methylation under varying conditions. (A) Schematic of Methylation Experiment 1. Wild-type *S.* Typhimurium (Strain 14028s) and an isogenic Δ*metJ* strain were cultured in LB or *Salmonella* Pathogenicity Island-2 (SPI-2)-inducing media and DNA was collected for SMRT-sequencing. Bacteria grown under identical conditions were harvested for RNA-sequencing. (B) Total number of m^{6}A bases observed across conditions does not dramatically change. (C) Analysis of motifs methylated reveals only the total number of ATGCA*T sites changes dramatically across conditions. For Panels B and C, bases were only included in the analysis if the base could confidently be called methylated or unmethylated across the four conditions tested, as well as the four control conditions in Supplemental Figure 1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

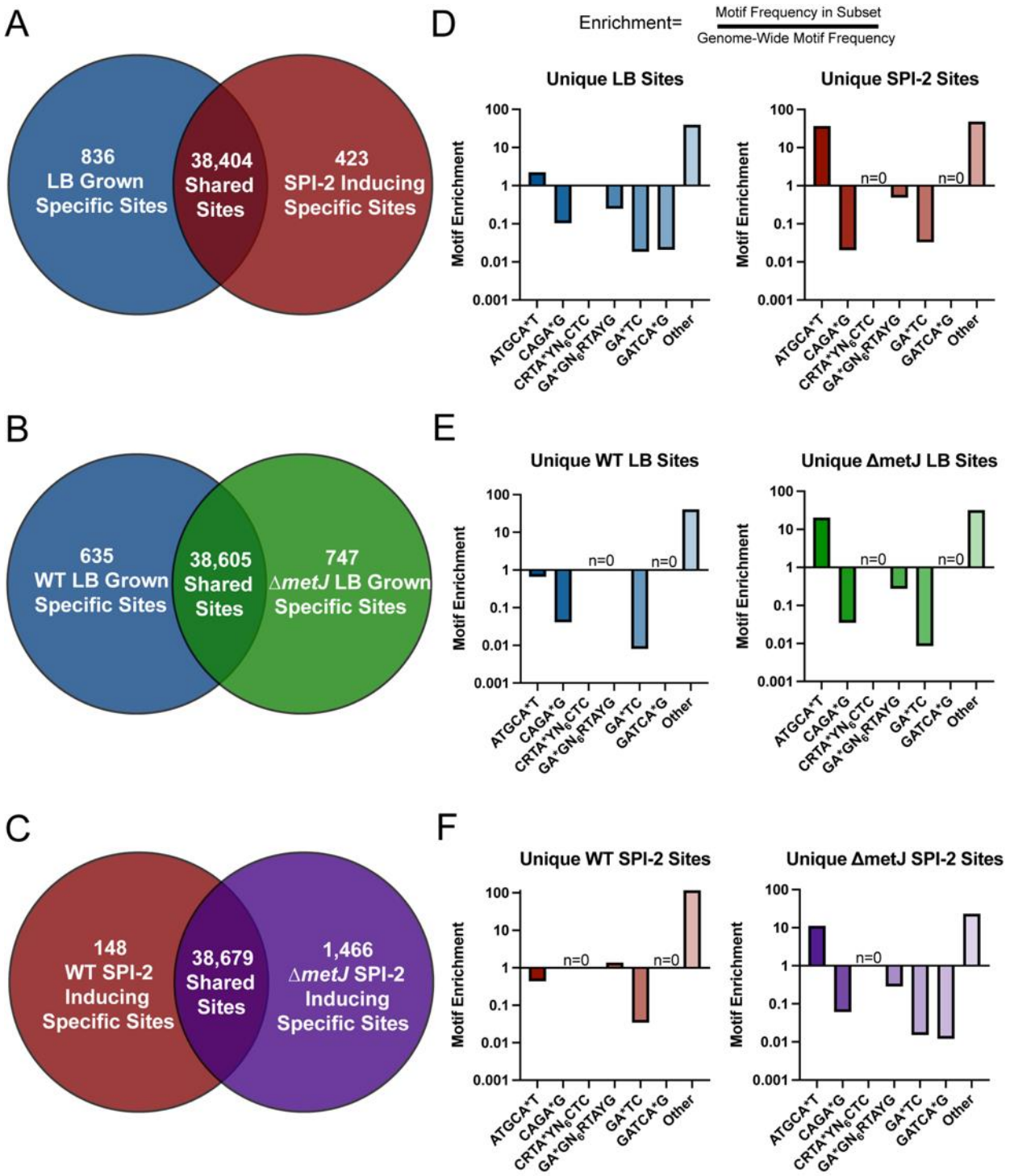
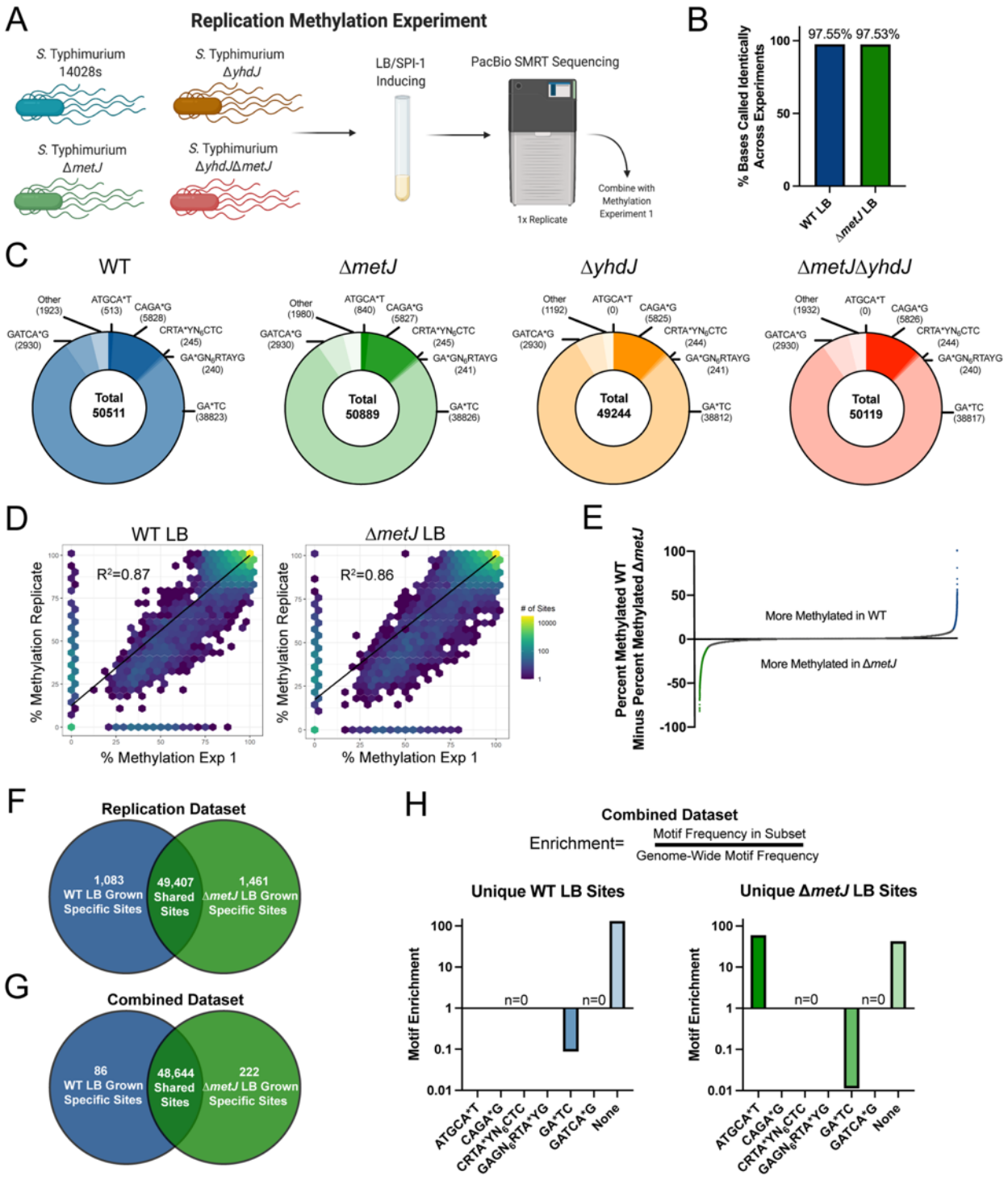


Figure 2: Analysis of subtle changes to the *S.* Typhimurium m^{6}A methylome in response to changing conditions reveals YhdJ as a dynamic methylase. (A-C) Quantification of shared and unique methylated sites between LB and SPI-2 inducing media (A), WT and Δ*metJ* bacteria grown in LB (B), and WT and Δ*metJ* bacteria grown in SPI-2 inducing media (C). (D-F) Identification of motifs enriched in methylation sites unique to each of the comparisons in Panels A-C. Motif enrichment was calculated by dividing the frequency of the motif among the uniquely methylated

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

bases by the genome-wide frequency of that motif within that condition (*ex.* For Panel D, frequency of ATGCA*T within unique WT SPI-2 sites = 242 ATGCA*T sites/423 unique SPI-2 sites (0.57); frequency of ATGCA*T within all WT SPI-2 = 600 ATGCA*T sites/38,843 detected motifs (0.015); enrichment = 0.57/0.015 = 37.04). For all panels, only bases that could be confidently called methylated or unmethylated in the four conditions here and in the four conditions in **Supplemental Figure 1** were considered.



bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 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-sequencing. (B) Approximately 97% of bases were called identically (methylated or unmethylated) in Methylation Experiment 1 and the Replicate Methylation Experiments. (C) Only ATGCA*T changes dramatically across tested conditions in the Replication Methylation Experiment. No ATGCA*T methylation was observed in Δ*yhdJ* mutants. (D) The observed Percent Methylation at each base is reproducible across 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. (E) Quantitative analysis reveals numerous sites are differentially methylated between wild-type and Δ*metJ*. Each dot represents the mean percent methylation in wild-type bacteria across the two experiments subtracted by the mean methylation in Δ*metJ* bacteria for each adenosine confidently called in both experiments. Blue and green dots mark bases where the mean difference is ≥10%. (F) Quantification of unique methylation sites in the Replication Experiment. For Panels C-F, bases were only included in the analysis if the base could confidently be called methylated or unmethylated across conditions. (G) Quantification of unique methylation sites in the Combined Dataset. (H) The ATGCA*T motif is enriched in uniquely methylated Δ*metJ* sites in the Combined Dataset. Motif enrichment was calculated by dividing the frequency of the motif among the uniquely methylated bases by the genome-wide frequency of that motif within that condition. For G and H, analyses were run on the "Combined Dataset," which includes bases that could confidently be called methylated or unmethylated for wild-type and Δ*metJ* bacteria grown in LB and were identically called in both experiments.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

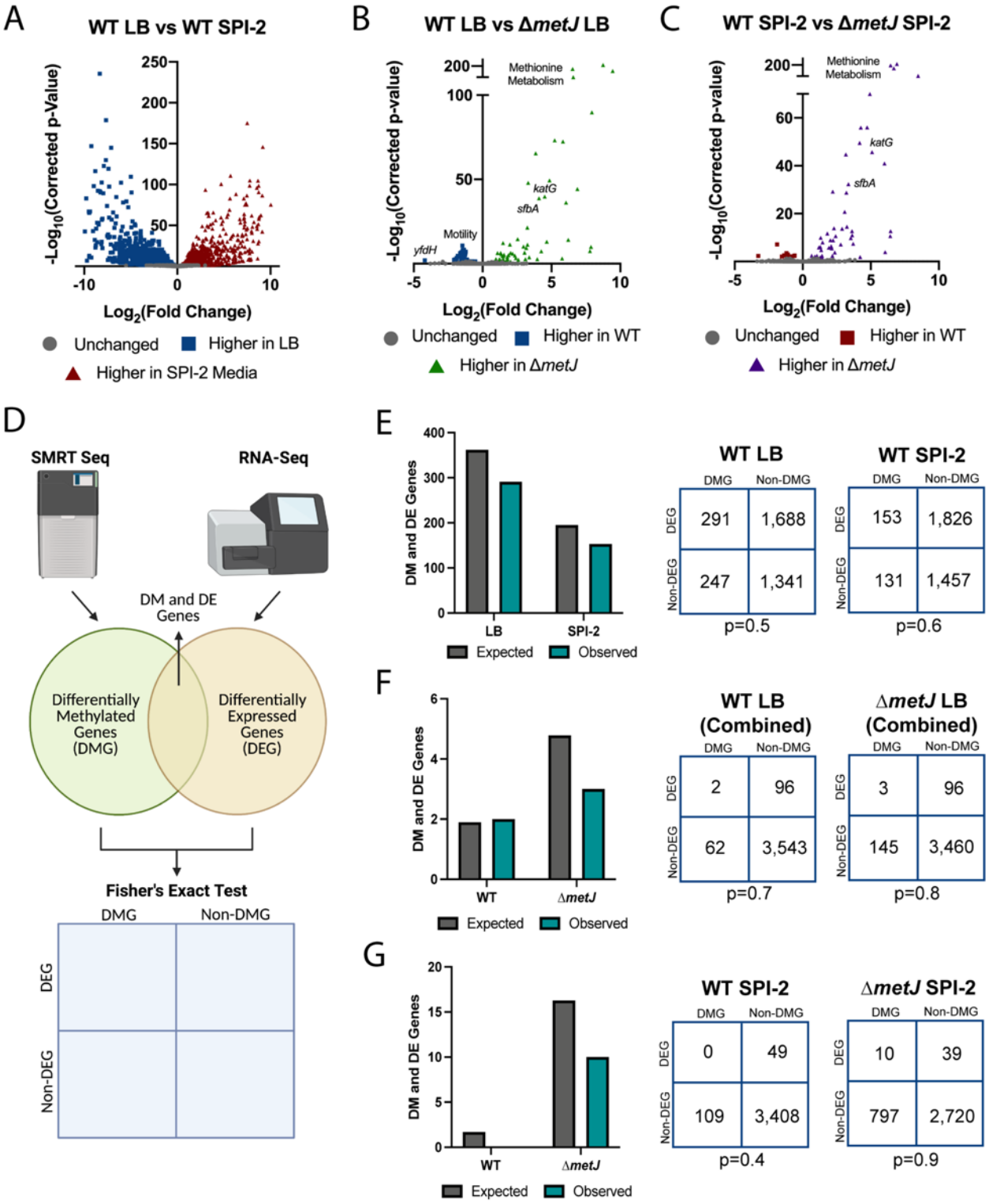
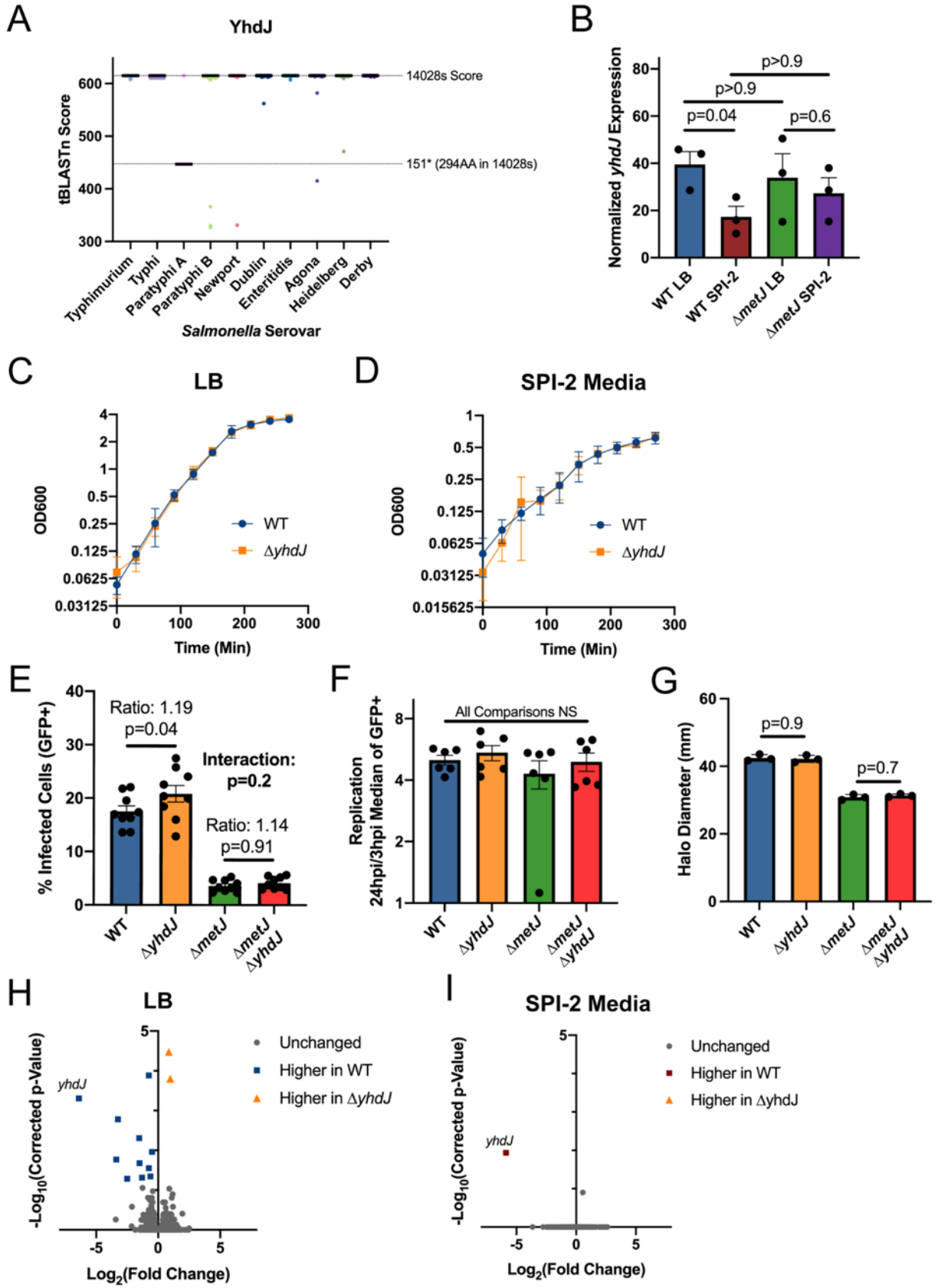


Figure 4: Differentially methylated genes are not enriched for transcriptomic changes. (A-C) RNA-seq reveals transcriptomic changes between LB grown and SPI-2 media grown wild-type bacteria (A), wild-type and Δ*metJ* bacteria grown in LB (B), and wild-type and Δ*metJ* bacteria grown in SPI-2 media (C). Corrected p-values generated by calculating the false discovery rate. (D) Schematic of RNA-seq and SMRT-seq integration. Each gene was determined to be differentially methylated (Differentially Methylated Gene, DMG), differentially expressed

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(Differentially Expressed Gene, DEG), differentially methylated and differentially expressed (DM and DE Gene), or neither differentially methylated nor expressed. Fisher's Exact Test was then used to determine whether there was an association between methylation and gene expression. (E-G) Differential methylation is not associated with differential expression. Observed and expected numbers of differentially methylation and differentially expressed genes were not significantly different when comparing uniquely methylated genes in LB vs SPI-2 media (E), wild-type vs Δ*metJ* in LB (F), or wild-type vs Δ*metJ* 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 number of differentially methylated genes in each category by the frequency of differentially expressed genes in that comparison. Numbers used for Fisher's Exact Test are shown on the right. Data for E and G used data from Methylation Experiment 1, F used the "Combined Dataset." For 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 5: YhdJ has limited impacts on *S*. Typhimurium biology under standard laboratory conditions. (A) 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 (67,68). Genomes were combined into a single FASTA file per serovar and blasted against the *S.* Typhimurium strain 14028s YhdJ protein sequence using BLAST+ (69). The BLAST score from the top 'n' hits were plotted, where 'n' is the number of genomes analyzed for that serovar. 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. (B) YhdJ reads obtained from the RNA-seq experiment in **Figure 4**. (C, 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 minutes. Error bars represent the standard error of the mean. (E, F) YhdJ is not required for *S.* Typhimurium uptake (E) or replication (F) in THP-1 monocytes. Cells were infected for 60 minutes with *S.* Typhimurium harboring an inducible-GFP plasmid before treatment with gentamicin. GFP was induced for 75 minutes before analysis by flow cytometry. Percent GFP+ and median of the GFP+ cells were measured three hours and twenty-four hours post infection. Panel E shows the amount of invasion that occurred by reporting the percent of infected cells at 3 hours post infection, and Panel F shows the replication that occurred following infection by dividing the median of the GFP+ cells at 24 hours post infection by the median of the GFP+ cells at 3 hours post infection. Data from 2-3 independent experiments, each dot represents an independent replicate, bars mark the mean, and error bars are the standard error 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 generated by two-way ANOVA with Šídák's multiple comparisons test. (G) YhdJ does not impact *S.* Typhimurium motility. Motility on soft agar was measured six hours after inoculating the agar and following migration at 37°C. Data are from three independent experiments and each dot is the average of 4-5 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 generated by one-way ANOVA with Šídák's multiple comparisons test. (H, I). YhdJ has limited impacts on the *S.* Typhimurium transcriptome in LB (H) and SPI-2 inducing media (I). Corrected p-values generated by calculating the false discovery rate.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

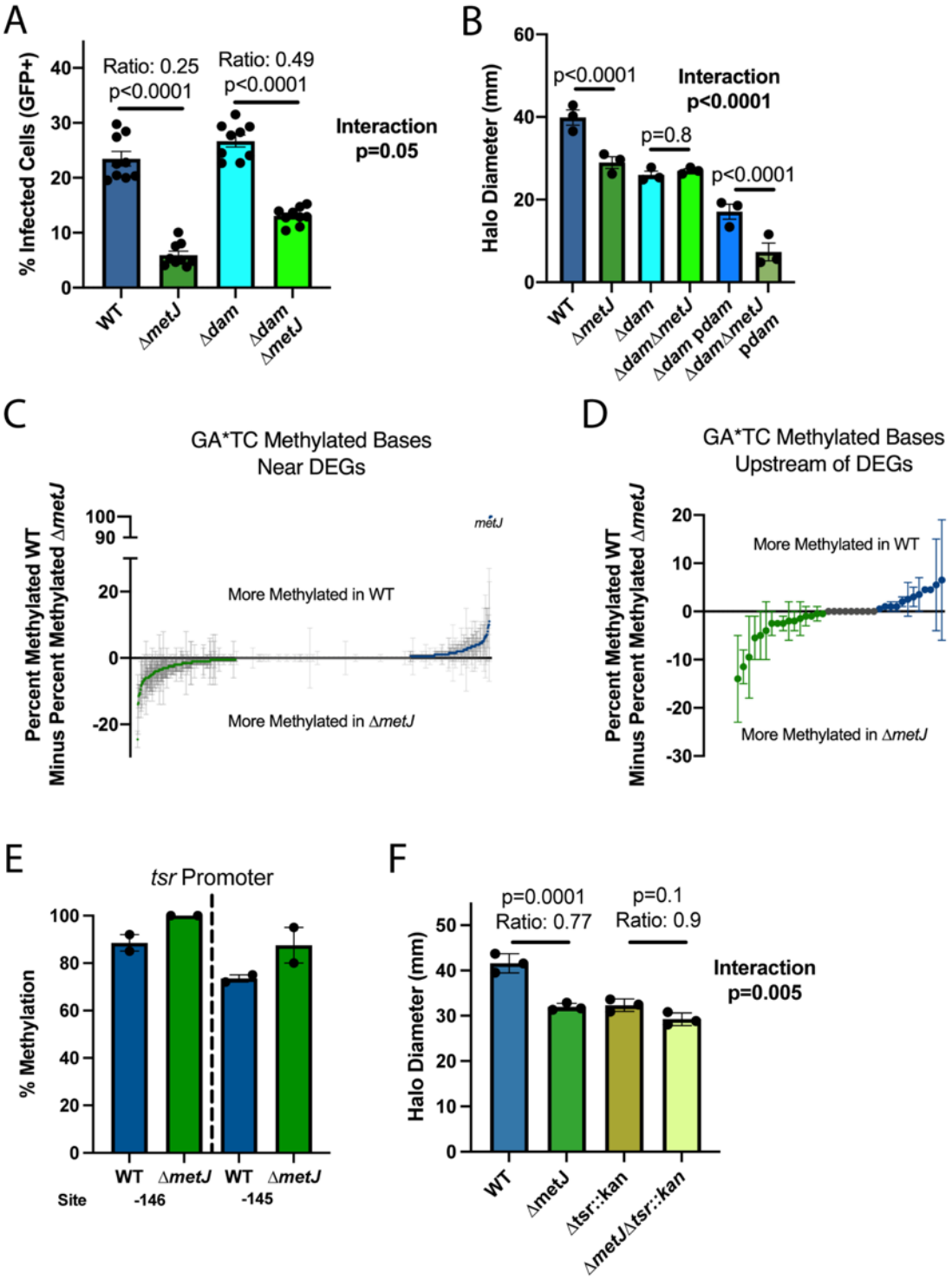


Figure 6: *dam* is epistatic to Δ*metJ* despite limited changes to the Δ*metJ* GA*TC methylome. (A) The impacts of Δ*metJ* on invasion partially depend on *dam*. THP-1 monocytes were infected for 60 minutes with *S.* Typhimurium harboring an inducible-GFP plasmid before treatment with gentamicin. GFP was induced for 75 minutes before analysis by flow cytometry. Percent GFP+ was measured three hours post infection. Data are from three experiments, each dot represents an independent replicate, the bars mark the mean, and the error bars are the standard error of the mean. (B) The impact of Δ*metJ* on motility depends entirely on *dam*. Motility on soft agar was measured six hours after inoculating the agar and following migration at 37°C. Data are from three independent experiments and each dot is the mean of 4-5 technical replicates, bars mark the mean, and error bars mark the standard error of the mean. (C,D)

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Quantitative analysis reveals subtle changes to the GA*TC methylome in Δ*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 Δ*metJ* bacteria grown in LB. 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 C and D, any base with greater than or less than 0 differential methylation is colored in green (more methylated in Δ*metJ*) or blue (more methylated in wild-type bacteria). (E) The *tsr* promoter is modestly hypermethylated in Δ*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 error of the mean. (F) The impacts of Δ*metJ* on motility are partially *tsr* dependent. Data are from three independent experiments and each dot is the mean of 4-5 technical replicates, bars mark the mean, and error bars mark the standard error 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 Šídák's multiple comparisons test.

Table 1: YhdJ does not enhance *S.* Typhimurium fitness in C57BL/6J mice

IP Infection (CFUs from Spleen)	**Δ*metJ*/WT**	**Δ*yhdJ*/WT**	**Δ*metJ*Δ*yhdJ*/WT**	**Δ*metJ*Δ*yhdJ*/Δ*metJ***
Number of Mice[$]	6	8	7	7
Median[#] *(95% Confidence Interval)*	**0.23** (0.07-0.47)	**1.31** (0.60-1.84)	**0.26** (0.13-0.78)	**1.09** (0.000065-2.19)
*p-value**	**0.002**	0.3	**0.002**	0.2
Oral Infection (CFUs from Spleen and Ileum)	**Δ*yhdJ*/WT (Spleen)**	**Δ*yhdJ*/WT (Ileum)**		
Number of Mice[$]	22	21		
Median[#] *(95% Confidence Interval)*	**1.65** (0.72-2.22)	**1.154** (0.74-1.55)		
*p-value**	**0.02**	0.4		

[$]All mice are age and sex matched, both sexes represented in all experiments, and all data are from at least two experiments
[#]Median competitive index value calculated by dividing the number colonies obtained of each genotype at four days post infection. Median > 1 = numerator strain outcompeted denominator.
*One-sample t-test on log transformed data

Table 2: Δ*yhdJ* Differential Gene Expression in LB

Gene ID	**Gene Name**	**Log_2Fold Change (Δ*yhdJ*/WT)**	**Corrected p-value**	**ATGCAT Motif? (Relative Location)**	**Methylated?**
STM14_4375	*dppA*	-0.75	2.95E-08	**Yes (Genic)**	**Yes**
STM14_4084	*yhdJ*	-6.39	2.25E-07	No	No
STM14_5353	*pyrB*	-3.25	1.12E-06	**Yes (Genic)**	No
STM14_4024	*codB*	-1.52	4.51E-06	No	No
STM14_0699	*cstA*	-0.48	1.25E-05	No	No

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

STM14_5352	*pyrI*	-3.38	2.32E-05	No	No
STM14_0819	*NA*	0.85	3.34E-05	No	No
STM14_4495	*pyrE*	-1.49	3.84E-05	No	No
STM14_5141	*acs*	-0.74	5.74E-05	No	No
STM14_1558	*yeaG*	-0.62	0.0001	No	No
STM14_3061	*uraA*	-1.3	0.0001	No	No
STM14_0078	*carB*	-2.49	0.0001	No	No
STM14_1885	*hdeB*	0.99	0.0002	No	No

Table 3: Unique GA*TC sites between wild-type and Δ*metJ S.* Typhimurium (Combined Dataset)

Unique WT LB GA*TC Sites					
Genomic Location	*Base*	*Genic Location*	*Gene*	*Product*	*Differentially Expressed?*
None	None	None	None	None	N/A
Unique Δ*metJ* LB GA*TC Sites					
Genomic Location	*Base*	*Genic Location*	*Gene*	*Product*	*Differentially Expressed?*
NC_016856.1	1177765	Upstream	STM14_1293	N-acetylneuraminate Epimerase	No
NC_016856.1	4289018	Downstream	STM14_4888	CDP-diacylglycerol Pyrophosphatase	No

Table 4: Percent methylation differences for GA*TC motifs upstream of Δ*metJ* differentially expressed genes following growth in LB

Base	**Closest Gene (STM14 Number)**	**Closest Gene (Common Name)**	**% Methylation (WT-Δ*metJ*) Methylation Experiment 1**	**% Methylation (WT-Δ*metJ*) Replication Experiment**	**% Methylation (WT-Δ*metJ*) Average**	**Gene Expression - Log_2Fold Change (Δ*metJ*/WT)**	**Gene Expression - Corrected p-value**
4802770	STM14_5446	*tsr*	-23	-5	**-14**	-1.28	0.006
4802769	STM14_5446	*tsr*	-15	-8	**-11.5**	-1.28	0.006
2033749	STM14_2341	*flhD*	-1	-18	**-9.5**	-0.53	0.05
3243659	STM14_3699	*serA*	-10	-1	**-5.5**	1.48	5.09E-13
4415526	STM14_5029	*aceB*	-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	*aceB*	-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	*aceB*	-3	1	**-1**	1.88	6.67E-06
4802790	STM14_5446	*tsr*	-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	*hmpA*	0	0	**0**	1.79	7.28E-05
2746735	STM14_3135	*hmpA*	0	0	**0**	1.79	7.27E-05
3243658	STM14_3699	*serA*	0	0	**0**	1.48	5.09E-13
3402626	STM14_3894		0	0	**0**	-1.23	7.87E-08

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.11.468322; this version posted November 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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	*aceB*	1	0	**0.5**	1.88	6.67E-06
1049501	STM14_1130	*ompF*	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	*tsr*	7	0	**3.5**	-1.28	0.006
1049502	STM14_1130	*ompF*	5	4	**4.5**	1.08	4.46E-07
3402625	STM14_3894		4	5	**4.5**	-1.23	7.86E-08
2072703	STM14_2394	*fliJ*	15	-4	**5.5**	-1.55	0.006
2072704	STM14_2394	*fliJ*	19	-6	**6.5**	-1.55	0.006