NAR Breakthrough Article

Contribution of DNA adenine methylation to gene expression heterogeneity in *Salmonella enterica*

María A. Sánchez-Romero*, **David R. Olivenza, Gabriel Gutiérrez and Josep Casadesús** [ID]*

Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Apartado 1095, Seville 41080, Spain

ABSTRACT

Expression of *Salmonella enterica* loci harboring undermethylated GATC sites at promoters or regulatory regions was monitored by single cell analysis. Cell-to-cell differences in expression were detected in ten such loci (*carA*, *dgoR*, *holA*, *nanA*, *ssaN*, *STM1290*, *STM3276*, *STM5308*, *gtr* and *opvAB*), with concomitant formation of ON and OFF subpopulations. The ON and OFF subpopulation sizes varied depending on the growth conditions, suggesting that the population structure can be modulated by environmental control. All the loci under study except *STM5308* displayed altered patterns of expression in strains lacking or overproducing Dam methylase, thereby confirming control by Dam methylation. Bioinformatic analysis identified potential binding sites for transcription factors OxyR, CRP and Fur, and analysis of expression in mutant backgrounds confirmed transcriptional control by one or more of such factors. Surveys of gene expression in pairwise combinations of Dam methylation-dependent loci revealed independent switching, thus predicting the formation of a high number of cell variants. This study expands the list of *S. enterica* loci under transcriptional control by Dam methylation, and underscores the relevance of the DNA adenine methylome as a source of phenotypic heterogeneity.

INTRODUCTION

Non mutational variation plays multiple roles in bacterial populations. For instance, formation of cell types with different surface antigens permits evasion of the immune system inside animal hosts ([1]). In biofilms, cell-to-cell variation can foster division of labor ([2]). Phenotypic heterogeneity can also facilitate survival in the presence of antibacterial agents ([3,4]) and of bacteriophages ([5,6]). The benefits of phenotypic variation remain hypothetical in other cases, mainly due to experimental difficulties. However, game theory analysis supports the view that cell-to-cell variation has adaptive value for bacterial populations, especially in variable and hostile environments ([7,8]). The variety of mechanisms that produce cell diversification, which is indicative of independent origin, provides additional, indirect evidence that phenotypic variation may have evolved by natural selection ([9–11]).

A widespread mechanism for formation of phenotypic variants in a population of isogenic bacterial cells is epigenetic control by DNA methylation ([12]). For instance, certain DNA methyltransferases belonging to restriction-modification systems undergo reversible switching between active and inactive states, thus generating cells in which the genome is methylated and cells in which the genome is nonmethylated ([13–15]). If the methylation state of promoters or regulatory regions affects transcription, each cell lineage displays a distinct transcriptomic profile. In turn, transcriptome differences can cause phenotypic differences, and the literature contains an increasing list of examples in which such differences affect virulence of bacterial pathogens ([13,14,16]).

Formation of bacterial lineages under DNA methylation control is also found in bacterial species that harbor DNA methyltransferases that are not part of restriction-modification systems ([11]). Paradigms of this class are the Dam methylase of gamma-proteobacteria and the CcrM methylase of alpha-proteobacteria, which methylate the adenosine moiety in 5′GATC3′ and 5′GANTC3′ motifs, respectively ([17–19]). In *Escherichia coli* and *Salmonella enterica*, most GATC sites are fully methylated except for a short period of hemimethylation after passage of the DNA replication fork ([17,18,20]). Certain GATC sites, however, can remain stably undermethylated (hemimethylated or nonmethylated) if the activity of Dam methylase is prevented by binding of proteins to cognate DNA sites ([12,18]). Hin-

*To whom correspondence should be addressed. Tel: +34 95 455 7105; Email: casadesus@us.es
Correspondence may also be addressed to María A. Sánchez-Romero. Email: mtsanchez@us.es

drance of Dam methylase activity is favored at GATC sites flanked by nucleotide sequences that decrease Dam processivity (21), and nonmethylation occurs when DNA methylase activity is blocked during two consecutive DNA replication rounds (22).

Combinations of methylated and undermethylated GATCs at regulatory regions upstream of promoters can regulate transcription in a bistable manner, giving rise to OFF and ON cells (22). In such systems, the existence of two or more binding sites for a transcription factor permits alternative binding patterns, one of which activates transcription. A paradigm of bistable expression regulated by Dam methylation is *pap,* a fimbrial operon of uropathogenic *E. coli* under transcriptional control the leucine-responsive regulatory protein, Lrp (23–25). Alternative patterns of Lrp binding and GATC undermethylation generate the OFF and ON states, which give rise to nonfimbriated and fimbriated subpopulations (23). Additional examples of formation of DNA methylation patterns involve the transcriptional regulators OxyR, Fur and HdfR (26–31). The number of GATC sites involved and the architecture of the regulatory region is different in each case but all have in common the formation of methylated and undermethylated GATCs (12).

Undermethylation of specific GATC sites can be detected by genome digestion with restriction enzymes followed by deep sequencing (32,33). An alternative procedure is whole genome analysis by single molecule real-time sequencing (SMRT), a technology that permits detection of N^6-methyl-adenine at a single base resolution (34). In this study, we have used previously published SMRT sequencing data (28) to assemble a catalog of undermethylated GATC sites in the genome of *S. enterica* ser. Typhimurium strain ATCC 14028. Single cell analysis using fluorescent reporter proteins GFP, mCherry and mOrange reveals that a fraction of such loci undergo heterogeneous expression under Dam methylation control, with concomitant formation of cells in ON and OFF transcriptional states. We thus identify novel loci with heterogeneous expression controlled by Dam methylation in *S. enterica*. We also show that seven such loci undergo ON/OFF switching in an independent manner, an observation that outlines the relevance of the DNA adenine methylome as a powerhouse for the formation of phenotypic variants.

MATERIALS AND METHODS

Bacterial strains, bacteriophages and strain construction

The strains of *S. enterica* used in this study belong to serovar Typhimurium and derive from strain ATCC 14028 (Supplementary Table S1). For simplicity, *S. enterica* serovar Typhimurium is named *S. enterica* throughout the text. Plasmid pTP166 (35) was kindly provided by Martin G. Marinus, University of Massachusetts, Worcester, MA. For construction of *gfp* transcriptional fusions, a fragment containing the promoterless green fluorescent protein (*gfp*) gene and the chloramphenicol resistance cassette was amplified from pZEP07 (36) using primers listed in Supplementary Table S2. The constructs were integrated into the chromosome of *S. enterica* using the lambda Red recombination

system (37). For construction of strains harboring mCherry transcriptional fusions, a DNA fragment containing the promoterless *mCherry* gene and the kanamycin resistance cassette was PCR-amplified from pDOC-R, an *mCherry*-containing derivative of plasmid pDOC (a gift from Steve Busby's lab, University of Birmingham, UK). For construction of strains harboring mOrange transcriptional fusions, the *mOrange2* gene and the kanamycin resistance cassette were PCR-amplified from pGEMT-mOrange (kindly provided by Carmen R. Beuzón, University of Málaga, Spain). Transductional crosses using phage P22 HT 105/1 [(38) and G. Roberts, unpublished] were used for construction of strains lacking individual transcription factors (OxyR, CRP, Lrp or Fur). Strains carrying pairwise combinations of GFP and either mCherry or mOrange fusions were also constructed by P22 HT transduction. The transduction protocol has been described elsewhere (39). To obtain phage-free isolates, transductants were purified by streaking on green plates (40).

Media and culture conditions

Bertani's lysogeny broth (LB) contained tryptone 10 g/l, yeast extract 5 g/l and NaCl 5 g/l. The composition of ISM (intracellular salts medium) was 5 mM KCl, 7.5 mM $(NH_4)_2SO_4$, 0.5 mM K_2SO_4, 38 mM glicerol (0.3% v/v), 0.1% casamino acids, 8 μM $MgCl_2$, 337 μM H_3PO_4, 80 mM MES (for titration to pH 5.8). Solid LB contained agar at 1.5% final concentration. All cultures were grown at 37°C. Oxygen limitation was achieved by growth without shaking in borosilicate tubes containining 5 ml of LB. To grow *oxyR* mutants on LB agar, 75 μl of a 10 mg/ml catalase solution (Sigma-Aldrich, St Louis, MO, USA) were spread on the surface of the plates.

Analysis of SMRT sequencing data

Bioinformatic analysis was performed on raw SMRT sequencing data described in a previous study (28). The Pacific Biosciences' SMRT Portal platform, v. 2.1.0 was used to identify modQVs at each position. These values were computed as the −10 log (*P*-value) based on the distributions of the kinetics of interpulse durations (IPD ratios). A modQV score of 20 is the minimum default threshold and corresponds to a *P*-value of 0.01 (28). All the undermethylated positions detected in this work are under this threshold.

Identification of undermethylated GATC sites in the genome of *Salmonella enterica* Serovar Typhimurium ATCC 14028

A homemade Perl script was used to identify undermethylated GATC sites among the 19 229 GATCs present in the *S. enterica* ATCC 14028 genome. Another in-house Perl script identified pairs or higher order clusters of GATC motifs separated by <256 nucleotides (41). The Kolmogorov–Smirnov test for two samples, implemented in PAST (42), was used to test whether the distribution of undermethylated GATCs in the *S. enterica* genome was homogeneous or heterogeneous.

Identification of DNA sequence motifs for binding of transcription factors

The FIMO tool from the MEME software suite (43) with default settings was used to scan a sequence database set of DNA sequences from nine loci (*carA, dgoR, holA, nanA, ssaN, STM1290, STM3276, gtr* and *opvAB*) containing 600 bp upstream and 100 pb downstream of the transcription start site. Individual matches to OxyR, CRP and Fur binding motifs were sought.

Flow cytometry

Flow cytometry was used to monitor expression of transcriptional GFP fusions. Data acquisition was performed using a Cytomics FC500-MPL cytometer (Beckman Coulter, Brea, CA, USA) and data were analyzed with FlowJo X version 10.0.7r software (Tree Star, Inc., Ashland, OR). *S. enterica* cultures were washed and re-suspended in phosphate-buffered saline (PBS) for fluorescence measurement. Fluorescence values for 100 000 events were compared with the data from the reporter-less control strain, thus yielding the fraction of ON and OFF cells.

Fluorescence microscopy

Strains containing individual or pairwise combinations of GFP and mCherry/mOrange fusions were grown at 37°C. Samples of 1.5 ml were collected by centrifugation at 3400 × g for 5 min. Cells were placed on an agarose slab (0.9% agarose/1% LB medium) pre-warmed at 37°C. Images were captured with a Zeiss Apotome fluorescence microscope equipped with a 100× Plan Apochromat objective and an incubation system that allows cultivation and observation of living cells. Pictures were taken using an Axiocam 506 camera, and the images were analyzed using ImageJ software (Wayne Rasband, Research Services Branch, National Institute of Mental Health, MD, USA).

RESULTS

Distribution of undermethylated GATC sites in the *Salmonella enterica* genome

The methylation state of GATC sites in the chromosome of *S. enterica* strain ATCC 14028 grown in LB was inferred from raw SMRT data obtained in a previous study (28). Bioinformatic analysis revealed the presence of 46 undermethylated GATC sites. Among such sites, 20 were hemimethylated, and 13 were nonmethylated (Supplementary Table S3). A Kolmogorov–Smirnov test for uniform distribution indicated that the 33 undermethylated GATCs were not randomly distributed in the *S. enterica* genome. A list of the 46 nonmethylated sites detected (20 in hemimethylated GATCs and 26 (13 × 2) in nonmethylated GATCs is presented in Supplementary Table S3, indicating the DNA strand and the genome location. Most undermethylated sites (42/46) were located in clusters that included 2, 3, or 4 GATC sites at distances below the 1/256 distance predicted for random occurrence (41). Table 1 includes only the 16 loci containing GATC sites at promoters or regulatory regions (known or putative), and indicates their methylation

state. The distribution of these sites was as follows: one undermethylated GATC in the *dgoR* gene; two in *carA, ssaN, slrA, STM1290* and *yihU*; three in *STM5047, STM5308, STM3726* and *STM2047*; four in *ftnB, gtr, opvAB* and *STM4889*; five in *holA*; and eight in *nanA*. The list includes undermethylated GATC sites previously shown to be involved in transcriptional control of the *gtr and opvAB operons* (28,29). This observation served as an internal control, indicating that the bioinformatic analysis was reliable.

Expression patterns of loci harboring undermethylated GATC sites in putative regulatory regions

The expression patterns of the 16 loci harboring undermethylated GATC sites potentially involved in transcriptional control (*dgoR, carA, ssaN, slrA, yihU, STM1290, STM5047, STM5308, STM3726, STM2047, nanA, ftnB opvAB, gtr, holA* and *STM4889*) were investigated using transcriptional fusions with the green fluorescent protein (GFP) gene. Each fusion was constructed downstream of the locus under study, and expression was monitored by flow cytometry. Cultures were grown in LB under aerobiosis, LB under microaerophilia and intracellular salts medium (ISM). These choices took into consideration the *Salmonella* lifestyle: LB under microaerophilia permits reductionist imitation of the intestinal environment (44) while ISM mimics the intracellular environment (45). Expression of six loci (*ftnB, slrA, yihU, STM2047, STM4889* and S*TM5047*) was low or absent under every growth condition (Supplementary Figure S1), and the loci were excluded from further analysis. GFP fusions in the remaining 10 loci were found to be active, and a remarkable observation was that expression appeared to be heterogeneous in one or more culture conditions, with formation of cells in OFF and ON states (Figure 1, panel A). Visualization of individual *S. enterica* cells by fluorescence microscopy confirmed the existence of ON and OFF cells (Figure 1, panel B).

For certain loci, different proportions of OFF and ON cells were found depending on the growth conditions, indicating that expression of such loci is susceptible to environmental influence. For instance, the OFF subpopulation found in LB under microaerophilia for *dgoR, carA* and *nanA* decreased or even disappeared in ISM and in LB under aerobiosis (Figure 1). Differences in subpopulation sizes depending on growth conditions were also detected for *ssaN, holA* and *STM1290* (Figure 1).

Role of Dam methylation in transcriptional control of undermethylated loci

Loci showing heterogeneous expression (*carA, dgoR, gtrA, holA, nanA, opvAB, ssaN, STM1290, STM3276* and *STM5308*) were tested for Dam methylation control by monitoring the levels of GFP fluorescence in three genetic backgrounds: the wild type strain, a *dam* mutant, and a strain carrying the *dam* gene on the multicopy plasmid pTP144 (35). GFP fluorescence was measured by flow cytometry under the three growth conditions described above. Both heat maps and histograms are shown because each of these representations is more informative depending on the locus (Supplementary Figure S2). A simplified version of

Table 1. Loci harboring undermethylated GATCs in regulatory regions

Locus	Gene product	Number of GATCs	Number of undermethylated GATCs	Position(s) of GATC undermethylated A(s)[a]	Methylation state[b]
carA	Carbamoyl-phosphate synthase small chain	2	2	−511, −206	N, H
dgoR	Galactonate operon transcriptional repressor	1	1	−147	H
ftnB	Ferritin-like protein	4	2	−340, −156	H, H
gtr	O-antigen glycotransferase	4	2	−69, −56	N, N
holA	DNA polymerase III, delta subunit	5	1	+4	H
nanA	N-acetylneuraminate lyase	8	1	−58	H
opvAB	O-antigen chain length regulation	4	2	−178, −105	N, N
slrA	Glucitol/sorbitol-specific enzyme IIC component	2	1	−86	H
ssaN	Type III secretion ATP synthase	2	1	−202	N
STM1290	N-acetylmannosamine-6-phosphate-2-epimerase	2	1	−465	N
STM2047	Hypothetical protein	3	1	−68	H
STM3726	Putative mannitol dehydrogenase	3	1	−68	H
STM4889	Putative Na$^+$/galactosidase symporter	4	1	−171	N
STM5047	Putative cytoplasmic protein	3	1	−108	H
STM5308	Sugar transporter	3	1	−66	N
yihU	Hypothetical oxidoreductase	2	1	−66	N

[a]Relative to known or predicted transcription start sites.
[b]N: nonmethylated; H: hemimethylated.

the same results is presented in Figure 2, which shows the expression pattern of each locus under the growth conditions that revealed Dam methylation-dependent control in a most conspicuous manner. *STM5308* is not included in Figure 2 because Dam-dependent regulation was not detected (see Supplementary Figure S2). The Dam methylation-dependent expression patterns observed for the remaining loci were as follows:

(i) In a *dam* background, the size of ON subpopulations decreased at four loci (*dgoR, STM1290, gtr* and *ssaN*), suggesting that GATC methylation is involved in transcriptional activation (Figure 2). In turn, a larger ON subpopulation was detected at two loci (*opvAB* and *holA*), suggesting Dam-dependent repression (Figure 2). Absence of Dam methylation had little effect, if any, on *carA, nanA* and *STM3726* expression (Figure 2). As above, *gtr* and *opvAB* served as internal controls as these loci have been previously shown to undergo transcriptional upregulation (*opvAB*) and transcriptional downregulation (*gtr*) in a *dam* background.

(ii) Overexpression of Dam methylase increased the size of CarA^{ON} and STM3726^{ON} subpopulations, thereby indicating Dam-dependent activation of transcription (Figure 2).

(iii) Detection of a smaller HolA^{ON} subpopulation upon Dam methylase overexpression confirmed Dam-dependent repression (Figure 2). At *nanA*, repression was especially severe, and Dam methylase overproduction abolished formation of the Nan^{ON} subpopulation (Figure 2).

Identification of regulators involved in transcriptional control

Undermethylation of GATC sites is usually a consequence of protein binding which hinders DNA methylase activity (11,12). The presence of undermethylated GATC sites

upstream of genes (Table 1 and Supplementary Table S3) thus suggested that these sites might be located within or near binding sites for transcription factors. Because transcription factors OxyR, LRP, CRP and Fur have been previously shown to direct undermethylated GATC formation (12), expression of the loci under study was monitored in the absence of individual transcription factors. For this purpose, *gfp* fusions were transferred to *oxyR, lrp, crp* and *fur* mutants using P22 HT transduction. The resulting strains were grown, as above, in LB under aerobiosis, LB under microaerophilia and ISM medium, and the expression of the *gfp* fusions was monitored by flow cytometry. Strains harboring the same fusions in a wild type background were used as controls. The results from these experiments are shown in Supplementary Figure S3. A simplified version of the same results is presented in Figure 3, showing the expression pattern of each locus under the growth conditions in which altered expression is most easily seen. With the exception of *STM3726*, the remaining loci under study changed their expression pattern in one or more mutant backgrounds (Figure 3). However, none of the loci under study was found to be under Lrp control. Changes in the expression patterns of *gtr* and *opvAB* in an *oxyR* background are in agreement with the literature (28,29) thus serving as internal controls for the trials (Figure 3). Relevant observations can be summarized as follows:

(i) The subpopulations of CarA^{OFF}, DgoR^{OFF}, NanA^{OFF}, SsaN^{OFF} and STM1290^{OFF} cells were larger in the absence of CRP, thereby suggesting transcriptional activation. In contrast, the size of the HolA^{ON} and OpvAB^{ON} subpopulations increased in a *crp* background, which may be indicative of repression by CRP. The ability of CRP to act either as a repressor or as an activator of transcription has been known for decades (46).
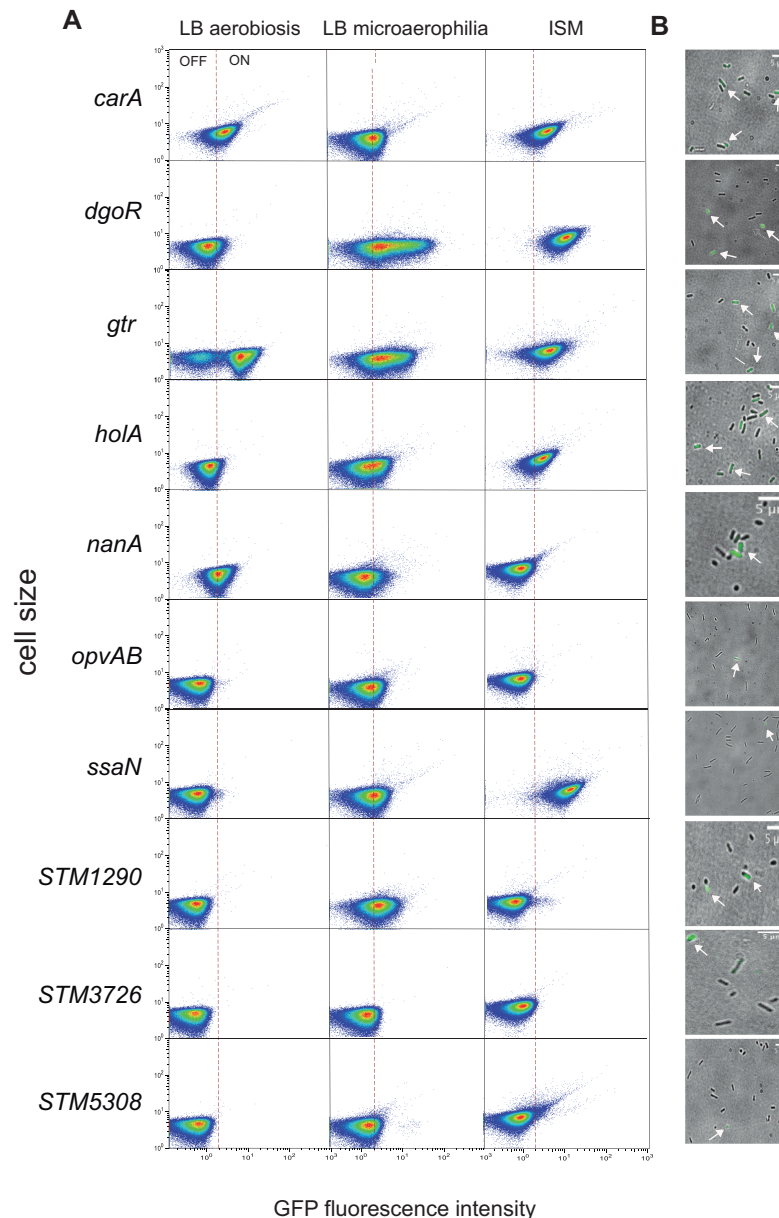
Figure 1. Expression of loci harboring undermethylated GATC sites at putative promoters or regulatory regions, monitored by single cell analysis. (**A**) Flow cytometry analysis of strains carrying *gfp* transcriptional fusions downstream of the loci under study. The GFP fluorescence intensity distribution was examined in cultures grown in LB under aerobiosis, LB under miare OKcroaerophilia and intracellular salts medium (ISM). Dot plots represent the forward scatter (cell size) versus fluorescence intensity. (**B**) Visualization of *Salmonella* cells by fluorescence microscopy. For each locus, the image shown was obtained under the growth condition in which bistability was most conspicuous (LB under microaerophilia for *carA*, *dgoR*, *gtr*, *holA*, *STM1290*, *STM3276* and *STM5308*, and LB under aerobiosis for *opvAB* and *ssaN*). White arrows signalize ON cells. Scale bars indicate 5 μm.

(ii) In the absence of Fur, the size of the OpvAB[ON] sub-population increased, suggesting Fur-mediated repression. A Dam methylation-dependent locus repressed by Fur has been described in *E. coli* (27,47).

(iii) In the absence of OxyR, the frequencies of SsaN[ON] and OvpAB[ON] cells decreased. The role of OxyR as an *opvAB* activator has been described previously (48).

To investigate whether the expression changes detected by flow cytometry might provide evidence for transcriptional control, potential DNA binding motifs for the transcription factors under study were sought at sequences flanking the GATCs. The FIMO tool from the MEME software suite was used for this purpose (49). Subsequently, Tomtom was used to compare the motifs found by MEME with known motifs contained at the Collect TF database, https://www.uniprot.org/database/DB-0198. Results from these searches are shown in Table 2 and Supplementary Table S4, and can be summarized as follows:

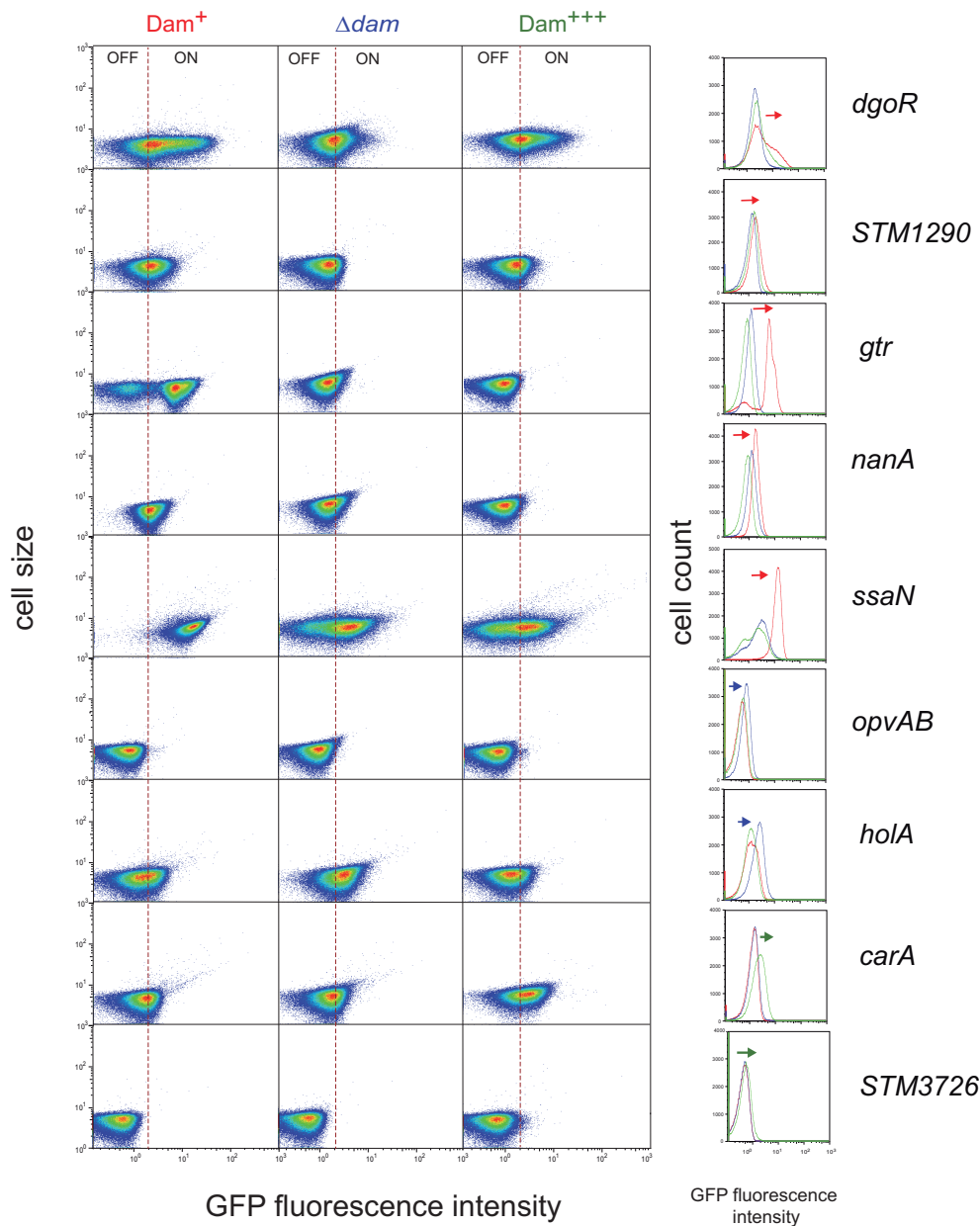(i) Putative CRP binding sites were identified upstream of *carA*, *dgoR*, *nanA, ssaN* and *STM1290* (Table 2),

Figure 2. Single cell analysis of Dam methylation-dependent control of *carA, dgoR, gtr, holA, nanA, opvAB, ssaN, STM1290* and *STM3726*. Gene expression was measured by flow cytometry in the wild type (left column and red line in the histograms), in a *dam* mutant (middle column and blue line in the histograms) and under Dam methylase overproduction (right column and green line in the histograms). Dot plots represent the forward scatter (cell size) versus fluorescence intensity. Histograms represent the frequency of cells with different GFP intensities, and changes in the pattern of expression are indicated with arrows. The growth condition shown for *dgoR, STM1290, holA* and *carA* is LB under microaerophilia; for *gtr, nanA, opvAB* and *STM3726* is LB under aerobiosis; and for *ssaN* is ISM minimal medium. More detailed information is provided in Supplementary Figure S2.

in agreement with the observation that these loci are under positive control by CRP. In *dgoR* and *nanA*, GATCs where found within the putative CRP binding sites while in *carA, STM1290* and *ssaN*, GATC sites were found near (but not within) the putative CRP binding sequence. Note that 'outsider', nearby GATC sites can anyway be involved in transcriptional control, presumably as a consequence of the formation of higher order DNA structures in regions bound by transcription factors active as dimers or multimers (28).

(ii) Putative OxyR binding sites were found in *gtr, opvAB*, and *ssaN*, which is consistent with genetic evidence indicating regulation by OxyR (Figure 3). For *gtr* and *opvAB*, regulation by OxyR is also in agreement with the literature (28,29,48).

(iii) Putative binding sites for Fur were detected upstream of both *opvAB* and *gtr* (Table 2). Fur-mediated regulation was indeed detected for *opvAB* (Figure 3) but not for *gtr*, perhaps because the experimental conditions were not appropriate.
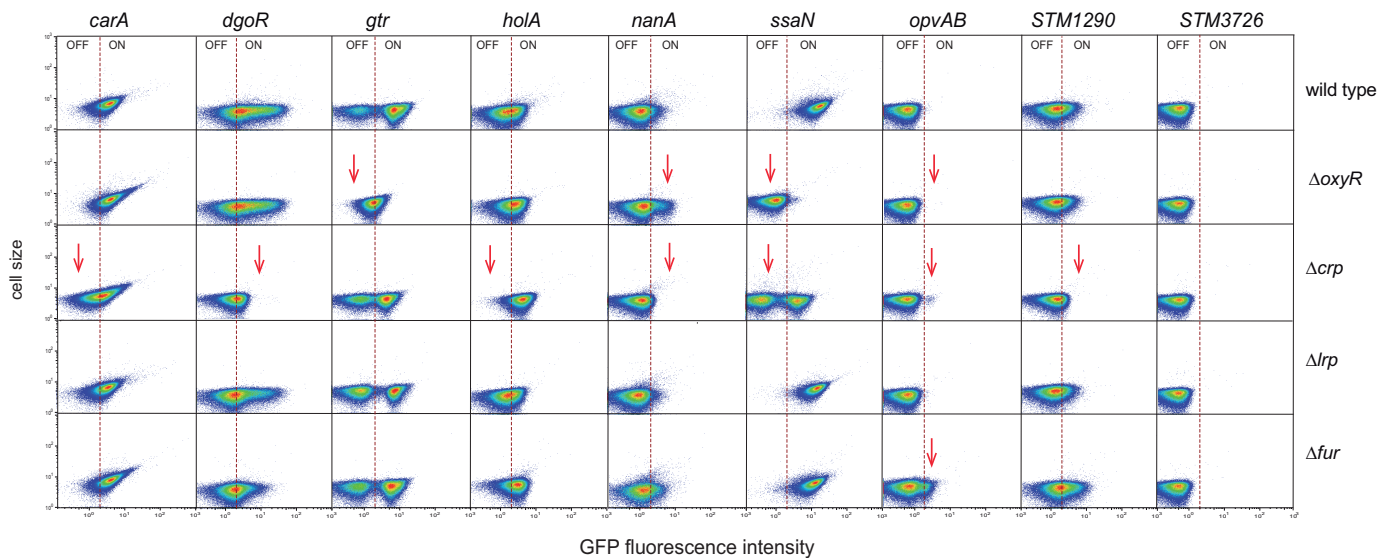
Figure 3. Flow cytometry analysis of *carA, dgoR, gtr, holA, nanA, opvAB, ssaN* and *STM1290* expression in the absence of individual transcriptional factors. Dot plots represent the forward scatter (cell size) versus fluorescence intensity. For *dgoR, holA, nanA, STM1290* and *STM3726*, the growth condition shown is LB under microaerophilia; for *gtr* and *opvAB* is LB under aerobiosis; and for *carA* and *ssaN* is ISM. More detailed information is provided in Supplementary Figure S3.

(iv) Failure to identify binding sites for CRP upstream of *holA* (Table 2) might be indicative of indirect control.

In summary, in most cases genetic evidence for regulation by a given transcription factor was accompanied by the detection of a putative binding site for such a factor. Because our trials have been restricted to transcription factors previously known to be involved in Dam-dependent transcription (12), one cannot exclude that expression of the loci under study may be subjected to additional controls.

A notion well established in the literature is that GATC undermethylation is typically caused by DNA binding proteins (e.g. transcription factors) that hinder Dam methyltransferase activity (12,24,50). On this ground, we examined the methylation state of the *nanA* promoter in *crp*+ and *crp*− backgrounds. Different methylation patterns were detected (Supplementary Figure S4), thus providing evidence that *nanA* GATC undermethylation is indeed a consequence of CRP binding. Methylation occlusion by transcription factors CRP, Fur and OxyR may likewise produce undermethylation at the other loci considered in this study.

Tests of co-ordinated *vs* independent switching

The existence of locus-specific expression patterns and the variety of transcription factors involved raised the possibility that each locus might undergo independent switching. However, certain regulators appeared to control more than one Dam-methylation locus, thus suggesting the alternative possibility that co-ordinated regulation might occur. To address this issue, tests were performed to determine whether the loci under study undergo ON−OFF switching in a co-ordinated manner or independently. For this purpose, strains that carried pairwise combinations of a GFP fusion and either an mCherry or mOrange fusion were constructed, and the presence of OFF and ON cells was monitored by fluorescence microscopy. Altogether, 10 pair-

wise combinations were tested: *ssaN* versus *gtr, ssaN* versus *carA, nanA* versus *gtr, nanA* versus *ssaN, opvAB* versus *gtr, opvAB* versus *ssaN, dgoR* versus *gtrA, dgoR* versus *ssaN, holA* versus *gtr* and *holA* versus *ssaN*. Cells that expressed either mCherry/mOrange or GFP were detected in all trials, together with cells that expressed both fusions and cells that did not express mCherry/mOrange nor GFP. These observations provide evidence that the loci under examination are capable of independent switching (Figure 4).

DISCUSSION

Heterogeneous gene expression in isogenic populations of cells can be the consequence of stochastic fluctuation of cell components, a phenomenon known as noise (51,52). In certain cases, cell-to-cell variations generated by noise can be propagated by feedback loops, giving rise to subpopulations of cells with distinct phenotypes (10,11,53). When two phenotypic lineages are formed, the phenomenon is known as bistability or bimodality (53,54). In bacterial populations, examples of noise-driven bistability include (among others) control of sporulation, activation of genetic competence, antibiotic resistance, biofilm formation and *myo*-inositol utilization (9,53,55).

Alongside bistable switches that combine random and deterministic components, bacteria can employ purely deterministic strategies to generate phenotypic lineages, and a well known mechanism is transcriptional control by DNA methylation (12). In gamma-proteobacteria, control of bistability by DNA adenine (Dam) methylation usually requires the formation of methylation patterns (combinations of methylated and undermethylated GATC sites) at promoters and/or nearby regions involved in transcriptional control. Undermethylation (stable hemimethylation or nonmethylation) of a GATC site occurs when Dam methyltransferase activity is hindered by a DNA-bound

Table 2. Identification of potential binding sites for transcription factors

Locus	Transcription factor involved	Predicted TF binding site (5′-3′)	Location relative to transcription start site
carA	CRP	CGATTGTAATTCTTATTACATTG	−499 to −477
dgoR	CRP	TTTTGTGATCTAAATTGTA	−154 to −132
gtr	CRP	TAACTTTAAACTATTGAATA	−100 to −82
	OxyR[b]	GATCGGTAACAACGATC	−134 to −119
	OxyR[b]	GATCGTTTATATCGATC	−70 to −55
	Fur	AATTAATAAGATAACAATAACTTTAAACTATTGAATA	−118 to −82
holA	CRP	Not found	
nanA	CRP	AAATGCGGTAGCGATCGCA	−145 to −127
	CRP	GTTTAGTGAAGCAGATCGCAC	−113 to −94
	OxyR	ATTGTGATTGTTGATGCCAT	−260 to −241
opvAB	CRP	AAAGTGTGCTCTGGTTTGCAA	−489 to −469
	OxyR[a]	ATAGTTTATATCTAT	−161 to −147
	OxyR[a]	GATCGATTTTAATTAT	−129 to −115
	OxyR[a]	ATAGTTTTTATCTAT	−88 to −74
	OxyR[a]	GATCGATATAACCAGT	−56 to −42
	Fur	CGTCACATAAAACAAAAACGATCAATTTTATTTATATGA	−199 to −161
ssaN	CRP	GTCTTATTTTCGGCACCTTAA	−435 to −415
	OxyR	CATTAATAAAATCGCTGAAACTTTGCAACGGCTTGT	−500 to −465
STM1290	CRP	ATGTGACCAAGCGCTAAAAAT	−428 to −408
	CRP	TGTGATGCATTTCATACTG	−283 to −265
	CRP	TATGACTGTGTTCACATTT	−102 to −84

[a]Sites described also by Cota *et al*. (28).
[b]Sites described also by Broadbent *et al*. (29).

protein (12). In *E. coli* and *Salmonella*, the list of proteins that are known to cause undermethylation of GATC sites includes transcription factors Lrp, OxyR, Fur and HdfR (12).

In this study, we have used undermethylated GATC sites in the *S. enterica* genome as indicators of transcriptional control by Dam methylation. Indeed, among 16 loci containing undermethylated GATCs at promoters or regulatory regions (known or putative), nine showed distinct expression patterns depending on Dam methylase availability (Figure 2). Detection of undermethylated DNA sites upon SMRT sequencing may thus be considered a fairly productive approach, alternative or complementary to genetic and transcriptomic screens (56,57), to identify genes under DNA methylation control.

A remarkable trait shared by the nine Dam-dependent loci under study is bistable expression, with concomitant formation of cells in OFF and ON transcriptional states (Figure 1). Altered OFF and ON subpopulation sizes upon lack and/or overproduction of Dam methyltransferase confirmed that bistability is under Dam methylation control (Figure 2 and Supplementary Figure S2). An additional observation was that the ON/OFF ratios for the Dam-dependent loci under study varied depending on the medium and/or the growth conditions (Figure 2 and Supplementary Figure S2), in agreement with seminal studies describing conditions that affected the formation of undermethylated GATC sites in the *E. coli* chromosome (58,59).

Evidence for environmental influence, together with the fact that undermethylation is usually caused by protein binding (18,50), led us to search for *S. enterica* factors that might be involved in transcriptional control of the loci under study. Single cell analysis in mutant strains lacking individual transcription factors OxyR, CRP and Fur confirmed that subpopulation formation was under the control of one

or more such factors (Figure 3 and Supplementary Figure S3). Detection of transcriptional control of *opvAB* and *gtr* by OxyR was confirmatory of published data (28,29,48), thus providing an internal control for the survey. Among the novel observations made, variations in the size of the OpvAB[ON] lineage were detected in the absence of either CRP or Fur (Figure 3 and Supplementary Figure S3), suggesting that these regulators may contribute to transcriptional control of the *opvAB* operon. Because *opvAB* controls O-antigen chain length (48), CRP and/or Fur might perhaps skew ON/OFF switching under specific environmental circumstances. Environmental control of bistability may be also envisaged, again in a speculative manner, for CRP-mediated control of *dgoR, nanA* and perhaps *holA*. Dam methylation-dependent control of *ssaN,* a gene of *Salmonella* pathogenicity island 2 (SPI-2) (60), might play a role in the interaction of *Salmonella* with the animal host. The *ssaN* gene product, an ATPase, is essential for the activity of the type III secretion apparatus encoded by SPI-2, which permits survival inside macrophages (61). Formation of SsaN[OFF] and SsaN[ON] subpopulations thus raises the possibility that SPI-2 expression inside macrophages may be bistable. Under laboratory conditions, formation of SsaN[OFF] and SsaN[ON] lineages is under positive control by OxyR and under negative control by CRP (Figure 3 and Supplementary Figure S3), which may suggest a complex regulation, perhaps responsive to environmental cues. Speculations on the physiological significance of Dam methylation-dependent bistability in housekeeping loci like *holA*, a gene involved in DNA replication, and *dgoR*, a metabolic gene, would be even more premature than for the genes discussed above.

In contrast with our tentative, uncertain speculations on the physiological roles of Dam-dependent bistability in the loci under study, a clear-cut conclusion was obtained in tri-
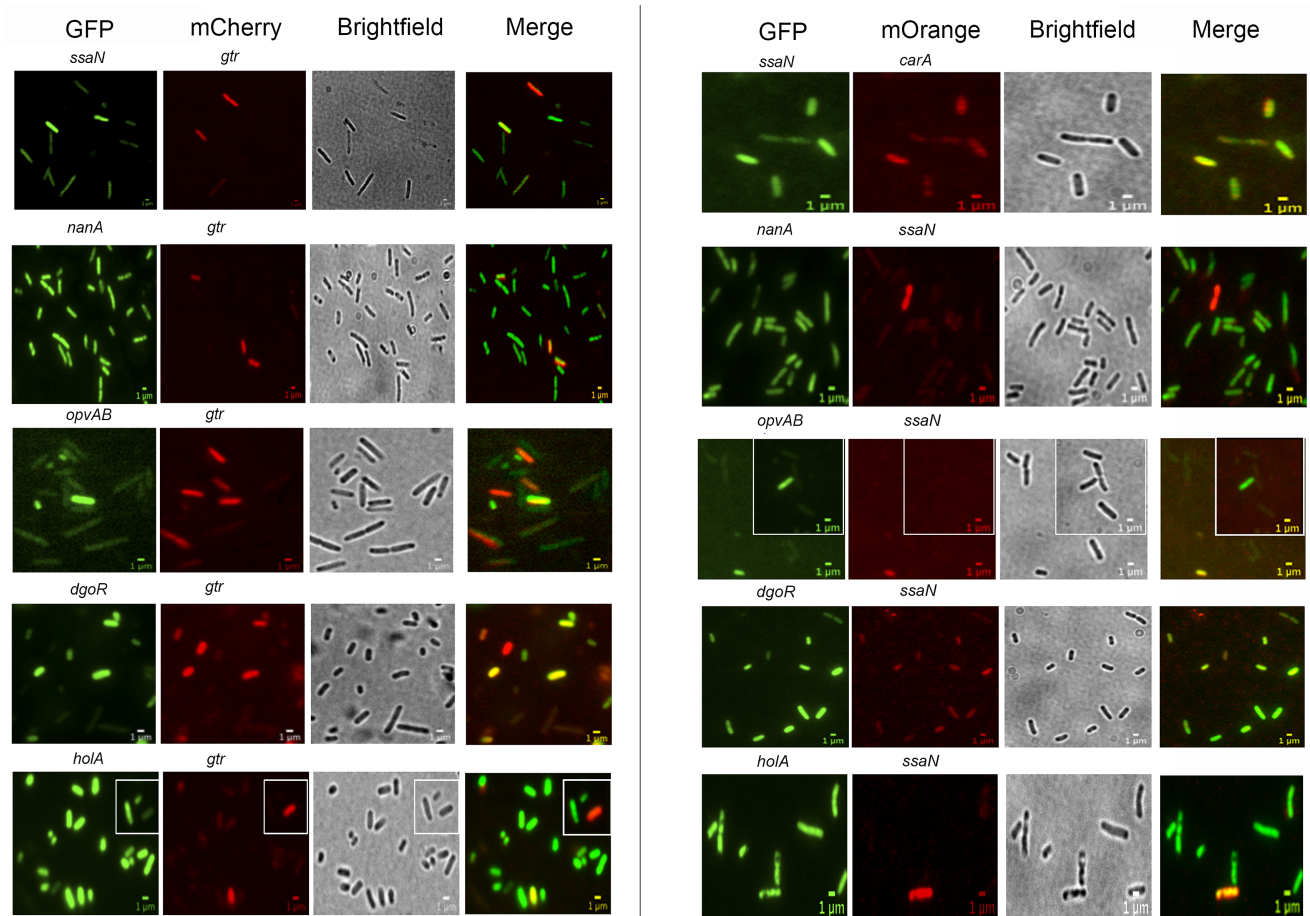
Figure 4. Tests of co-ordinated and/or independent switching of loci under transcriptional control by Dam methylation. Fluorescence images of *S. enterica* cells are shown, visualizing the expression of 10 pairwise combinations of gene fusions (*ssaN* versus *gtr*, *ssaN* versus *carA*, *nanA* versus *gtr*, *nanA* versus *ssaN*, *opvAB* versus *gtr*, *opvAB* versus *ssaN*, *dgoR* versus *gtrA*, *dgoR* versus *ssaN*, *holA* versus *gtr* and *holA* versus *ssaN*). Each row shows a strain containing a GFP and either an mCherry or an mOrange transcriptional fusion. The growth condition shown for *ssaN* versus *gtr*, *nanA* versus *gtr*, *opvAB* versus *gt*, *ssaN* versus *carA*, *nanA* versus *ssaN*, *opvAB* versus *ssaN*, *dgoR* versus *ssaN* and *holA* versus *ssaN* is LB under aerobiosis; for *dgoR* versus *gtr*, and *holA* versus *gtr*, the growth condition shown is LB under microaerophilia. Scale bars indicate 1 μm.

als aimed at determining whether switching of the loci under study was co-regulated or independent. Indeed, fluorescence microscopy observation of ON and OFF cells in strains that carried pairwise combinations of distinct fluorescent fusions revealed independent switching of seven loci (*ssaN, gtr, carA, dgoR, nanA, opvAB* and *holA*) (Figure 4). Dam-dependent control of bistability can thus generate a high degree of transcriptional heterogeneity, as illustrated by the simple calculation that independent switching of n bistable loci can produce 2^n types of cell variants. If this prediction is fulfilled in natural environments, independent switching of the seven loci shown in Figure 4 may be able to produce 128 types of cells. Transcription factors responsive to environmental conditions may further modulate phenotypic heterogeneity by controlling the subpopulation sizes.

Despite its relative success, our survey of the formation of phenotypic cell variants under Dam methylation control may fall short to ponder the contribution of DNA methylation to nonmutational polymorphism. After all, the identification of loci harboring undermethylated GATC sites and the analysis of their expression patterns were done under specific laboratory conditions. Furthermore, the Dam methylase is merely one example among the many DNA methyltransferases present in bacterial genomes (62,63). Undermethylated DNA methyltransferase targets are frequently detected in SMRT sequencing trials (62), thus raising the possibility that DNA methylation hindrance by transcription factors may be a common phenomenon (33). Furthermore, target undermethylation and concomitant lineage formation can be also produced by phase variation in DNA methyltransferase synthesis, a phenomenon that may be widespread among human pathogens (13–15,64). DNA methylation can thus be a powerhouse for the formation of phenotypic variants in bacterial populations, and the known examples of DNA methylation-dependent control of gene expression may be the tip of an iceberg whose magnitude remains unknown.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Finlay,B.B. and McFadden,G. (2006) Anti-immunology: evasion of the host immune system by bacterial and viral pathogens. *Cell*, **124**, 767–782.
2. van Gestel,J., Vlamakis,H. and Kolter,R. (2015) Division of labor in biofilms: the ecology of cell differentiation. *Microbiol. Spectr.*, **3**, MB-0002–2014.
3. Dewachter,L., Fauvart,M. and Michiels,J. (2019) Bacterial heterogeneity and antibiotic survival: understanding and combatting persistence and heteroresistance. *Mol. Cell*, **76**, 255–267.
4. Adam,M., Murali,B., Glenn,N.O. and Potter,S.S. (2008) Epigenetic inheritance based evolution of antibiotic resistance in bacteria. *BMC Evol. Biol.*, **8**, 52.
5. Cota,I., Sanchez-Romero,M.A., Hernandez,S.B., Pucciarelli,M.G., Garcia-Del Portillo,F. and Casadesus,J. (2015) Epigenetic control of Salmonella enterica O-Antigen chain length: A tradeoff between virulence and bacteriophage resistance. *PLoS Genet.*, **11**, e1005667.
6. Turkington,C.J.R., Morozov,A., Clokie,M.R.J. and Bayliss,C.D. (2019) Phage-resistant phase-variant sub-populations mediate herd immunity against bacteriophage invasion of bacterial meta-populations. *Front. Microbiol.*, **10**, 1473.
7. Kussell,E. and Leibler,S. (2005) Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, **309**, 2075–2078.
8. Wolf,D.M., Vazirani,V.V. and Arkin,A.P. (2005) A microbial modified prisoner's dilemma game: how frequency-dependent selection can lead to random phase variation. *J. Theor. Biol.*, **234**, 255–262.
9. Schroter,L. and Dersch,P. (2019) Phenotypic diversification of microbial pathogens – cooperating and preparing for the future. *J. Mol. Biol.*, **431**, 4645–4655.
10. Veening,J.W., Smits,W.K. and Kuipers,O.P. (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.*, **62**, 193–210.
11. Casadesus,J. and Low,D.A. (2013) Programmed heterogeneity: epigenetic mechanisms in bacteria. *J. Biol. Chem.*, **288**, 13929–13935.
12. Sanchez-Romero,M.A. and Casadesus,J. (2020) The bacterial epigenome. *Nat. Rev. Microbiol.*, **18**, 7–20.
13. Phillips,Z.N., Husna,A.U., Jennings,M.P., Seib,K.L. and Atack,J.M. (2019) Phasevarions of bacterial pathogens - phase-variable epigenetic regulators evolving from restriction-modification systems. *Microbiology*, **165**, 917–928.
14. De Ste Croix,M., Vacca,I., Kwun,M.J., Ralph,J.D., Bentley,S.D., Haigh,R., Croucher,N.J. and Oggioni,M.R. (2017) Phase-variable methylation and epigenetic regulation by type I restriction-modification systems. *FEMS Microbiol. Rev.*, **41**, S3–S15.
15. Vasu,K. and Nagaraja,V. (2013) Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol. Mol. Biol. Rev.*, **77**, 53–72.
16. Adhikari,S. and Curtis,P.D. (2016) DNA methyltransferases and epigenetic regulation in bacteria. *FEMS Microbiol. Rev.*, **40**, 575–591.
17. Marinus,M.G. (1996) Methylation of DNA. In: Neidhardt,F.C., Curtiss,R., Ingraham,J.L., Lin,E.C.C., Low,K.B., Magasanik,B., Reznikoff,W.S., Riley,M., Schaechter,M. and Umbarger,H.E. (eds). *Escherichia coli and Salmonella: Cellular and Molecular Biology*. ASM Press, Washington, D. C., pp. 782–791.
18. Wion,D. and Casadesus,J. (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat. Rev. Microbiol.*, **4**, 183–192.
19. Mouammine,A. and Collier,J. (2018) The impact of DNA methylation in Alphaproteobacteria. *Mol. Microbiol.*, **110**, 1–10.
20. Lobner-Olesen,A., Skovgaard,O. and Marinus,M.G. (2005) Dam methylation: coordinating cellular processes. *Curr. Opin. Microbiol.*, **8**, 154–160.
21. Peterson,S.N. and Reich,N.O. (2006) GATC flanking sequences regulate Dam activity: evidence for how Dam specificity may influence pap expression. *J. Mol. Biol.*, **355**, 459–472.
22. Low,D.A. and Casadesus,J. (2008) Clocks and switches: bacterial gene regulation by DNA adenine methylation. *Curr. Opin. Microbiol.*, **11**, 106–112.
23. van der Woude,M., Braaten,B. and Low,D. (1996) Epigenetic phase variation of the *pap* operon in *Escherichia coli. Trends Microbiol.*, **4**, 5–9.
24. Hernday,A., Krabbe,M., Braaten,B. and Low,D. (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **29**, 29.
25. Hernday,A., Braaten,B. and Low,D. (2004) The intricate workings of a bacterial epigenetic switch. *Adv. Exp. Med. Biol.*, **547**, 83–89.
26. Haagmans,W. and van der Woude,M. (2000) Phase variation of Ag43 in *Escherichia coli*: Dam-dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription. *Mol. Microbiol.*, **35**, 877–887.
27. 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.
28. Cota,I., Bunk,B., Sproer,C., Overmann,J., Konig,C. and Casadesus,J. (2016) OxyR-dependent formation of DNA methylation patterns in OpvAB^{OFF} and OpvAB^{ON} cell lineages of *Salmonella enterica. Nucleic Acids Res.*, **44**, 3595–3609.
29. 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.
30. Garcia-Pastor,L., Sanchez-Romero,M.A., Gutierrez,G., Puerta-Fernandez,E. and Casadesus,J. (2018) Formation of phenotypic lineages in *Salmonella enterica* by a pleiotropic fimbrial switch. *PLoS Genet.*, **14**, e1007677.
31. 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.
32. Dalia,A.B., Lazinski,D.W. and Camilli,A. (2013) Characterization of undermethylated sites in *Vibrio cholerae. J. Bacteriol.*, **195**, 2389–2399.
33. Ardissone,S., Redder,P., Russo,G., Frandi,A., Fumeaux,C., Patrignani,A., Schlapbach,R., Falquet,L. and Viollier,P.H. (2016) Cell cycle constraints and environmental control of local DNA hypomethylation in alpha-Proteobacteria. *PLos Genet.*, **12**, e1006499.
34. 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. *Nat. Methods*, **7**, 461–465.
35. Marinus,M.G., Poteete,A. and Arraj,J.A. (1984) Correlation of DNA adenine methylase activity with spontaneous mutability in *Escherichia coli* K-12. *Gene*, **28**, 123–125.
36. Hautefort,I., Proenca,M.J. and Hinton,J.C. (2003) Single-copy green fluorescent protein gene fusions allow accurate measurement of *Salmonella* gene expression in vitro and during infection of mammalian cells. *Appl. Environ. Microbiol.*, **69**, 7480–7491.
37. 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.
38. Schmieger,H. (1972) Phage P22-mutants with increased or decreased transduction abilities. *Mol. Gen. Genet.*, **119**, 75–88.

39. Garzon,A., Cano,D.A. and Casadesus,J. (1995) Role of Erf recombinase in P22-mediated plasmid transduction. *Genetics*, **140**, 427–434.

40. 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.

41. Zyskind,J.W. and Smith,D.W. (1980) Nucleotide sequence of the *Salmonella typhimurium* origin of DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 2460–2464.

42. Hammer,Ø., Harper,D.A.T. and Ryan,P.D. (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electron.*, **4**, 9.

43. Grant,C.E., Bailey,T.L. and Noble,W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017–1018.

44. Sanchez-Romero,M.A. and Casadesus,J. (2018) Contribution of SPI-1 bistability to *Salmonella enterica* cooperative virulence: insights from single cell analysis. *Sci. Rep.*, **8**, 14875.

45. Beuzon,C.R., Banks,G., Deiwick,J., Hensel,M. and Holden,D.W. (1999) pH-dependent secretion of SseB, a product of the SPI-2 type III secretion system of *Salmonella typhimurium*. *Mol. Microbiol.*, **33**, 806–816.

46. Saier,M.H. Jr and Ramseier,T.M. (1996) The catabolite repressor/activator (Cra) protein of enteric bacteria. *J. Bacteriol.*, **178**, 3411–3417.

47. Brunet,Y.R., Bernard,C.S. and Cascales,E. (2020) Fur-Dam regulatory interplay at an internal promoter of the enteroaggregative *Escherichia coli* Type VI Secretion *sci1* Gene Cluster. *J. Bacteriol.*, **202**, e00075-20.

48. 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.

49. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.

50. Casadesus,J. and Low,D. (2006) Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.*, **70**, 830–856.

51. Silva-Rocha,R. and de Lorenzo,V. (2010) Noise and robustness in prokaryotic regulatory networks. *Annu. Rev. Microbiol.*, **64**, 257–275.

52. Raser,J.M. and O'Shea,E.K. (2005) Noise in gene expression: origins, consequences, and control. *Science*, **309**, 2010–2013.

53. Dubnau,D. and Losick,R. (2006) Bistability in bacteria. *Mol. Microbiol.*, **61**, 564–572.

54. Graumann,P.L. (2006) Different genetic programmes within identical bacteria under identical conditions: the phenomenon of bistability greatly modifies our view on bacterial populations. *Mol. Microbiol.*, **61**, 560–563.

55. Garcia-Pastor,L., Puerta-Fernandez,E. and Casadesus,J. (2019) Bistability and phase variation in *Salmonella enterica*. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1862**, 752–758.

56. Torreblanca,J. and Casadesus,J. (1996) DNA adenine methylase mutants of *Salmonella typhimurium* and a novel dam- regulated locus. *Genetics*, **144**, 15–26.

57. 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. *J. Bacteriol.*, **188**, 8160–8168.

58. Ringquist,S. and Smith,C.L. (1992) The *Escherichia coli* chromosome contains specific, unmethylated *dam* and *dcm* sites. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 4539–4543.

59. Hale,W.B., van der Woude,M.W. and Low,D.A. (1994) Analysis of nonmethylated GATC sites in the *Escherichia coli* chromosome and identification of sites that are differentially methylated in response to environmental stimuli. *J. Bacteriol.*, **176**, 3438–3441.

60. Jennings,E., Thurston,T.L.M. and Holden,D.W. (2017) *Salmonella* SPI-2 type III secretion system effectors: Molecular mechanisms and physiological consequences. *Cell Host Microbe*, **22**, 217–231.

61. Yoshida,Y., Miki,T., Ono,S., Haneda,T., Ito,M. and Okada,N. (2014) Functional characterization of the type III secretion ATPase SsaN encoded by *Salmonella* pathogenicity island 2. *PLoS One*, **9**, e94347.

62. Blow,M.J., Clark,T.A., Daum,C.G., Deutschbauer,A.M., Fomenkov,A., Fries,R., Froula,J., Kang,D.D., Malmstrom,R.R., Morgan,R.D. *et al.* (2016) The epigenomic landscape of prokaryotes. *PLoS Genet.*, **12**, e1005854.

63. 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.

64. Srikhanta,Y.N., Fox,K.L. and Jennings,M.P. (2009) The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nat. Rev. Microbiol.*, **8**, 196–206.