

# Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*

Taku Oshima,<sup>1</sup> Chieko Wada,<sup>1,2</sup> Yuya Kawagoe,<sup>3</sup> Takeshi Ara,<sup>1,3</sup> Maki Maeda,<sup>1,2</sup> Yasushi Masuda,<sup>1</sup> Sota Hiraga<sup>4</sup> and Hirotada Mori<sup>1,5\*</sup>

<sup>1</sup>CREST, JST (Japan Science and Technology), Japan.

<sup>2</sup>Institute for Virus Research, Kyoto University, Kyoto 606-8507, Japan.

<sup>3</sup>Department of Cell Biology, Nara Institute of Science and Technology, Ikoma, 630-0101, Japan.

<sup>4</sup>Department of Molecular Cell Biology, Institute of Molecular Embryology and, Genetics, Kumamoto University School of Medicine, Kumamoto 862-0976, Japan.

<sup>5</sup>Research and Education Center for Genetic Information, Nara Institute of Science and Technology, Ikoma 630-0101, Japan.

## Summary

**Deoxyadenosine methyltransferase (Dam) methylates the deoxyadenine residues in 5'-GATC-3' sequences and is important in many cellular processes in *Escherichia coli*. We performed a computational analysis of the entire *E. coli* genome and confirmed that GATC sequences are distributed unevenly in regulatory regions, which suggests that Dam might regulate gene transcription. To test this, a high-density DNA microarray of 4097 *E. coli* genes was constructed and used to assess the gene expression profiles of the wild type and the *dam-16::kam* mutant strain grown under four different conditions. We also used two-dimensional electrophoretic analysis of the proteome to assess the protein profiles. The expression of a large number of genes was affected by the *dam* deficiency. Genes involved in aerobic respiration, stress and SOS responses, amino acid metabolism and nucleotide metabolism were expressed at higher levels in the mutant cells, especially in aerobic conditions. In contrast, transcription of genes participating in anaerobic respiration, flagella biosynthesis, chemotaxis and motility was decreased in the *dam* mutant strain under both aerobic and low aerobic**

**conditions. Thus, Dam-controlled genes are involved in adjusting the metabolic and respiratory pathways and bacterial motility to suit particular environmental conditions. The promoters of most of these Dam-controlled genes were also found to contain GATC sequences that overlap with recognition sites for two global regulators, fumarate nitrate reduction (Fnr) and catabolite activator protein (CRP). We propose that Dam-mediated methylation plays an important role in the global regulation of genes, particularly those with Fnr and CRP binding sites.**

## Introduction

The 4.6 Mbp *Escherichia coli* genome encodes about 4300 open reading frames (ORFs) (Blattner *et al.*, 1997), the functions of about 50% of which remain unknown. To understand the global gene regulation of the *E. coli* genome, its gene expression under various conditions has been comprehensively investigated using transcriptome and proteome analytical methods. This includes the transcriptome analyses using DNA microarrays that were performed to study the gene expression that takes place in response to changing environment conditions, during the heat shock response and as a result of gene disruption (Richmond *et al.*, 1999; Tao *et al.*, 1999; Arfin *et al.*, 2000). In addition, the proteome approach using two-dimensional gel electrophoresis has been performed with strains with mutations in nucleoid proteins such as IHF, H-NS or Fis (Nyström, 1995; Laurent-Winter *et al.*, 1997; Choe *et al.*, 1999), as well as to study the response to environmental stimuli (cold shock, heat shock and exogenous pyrophosphate) (Van Bogelen and Neidhardt, 1990; Biville *et al.*, 1996) and to assess the alterations that occur during the change in growth phase (Nyström *et al.*, 1996).

Dam (deoxyadenosine methyltransferase) methylates the adenine residue within 5'-GATC-3' sequences in double-stranded DNA. This methylation is known to play important physiological roles in *E. coli*. This is particularly demonstrated by Dam-defective mutants, which have highly pleiotropic changes including increased mutability, hyper-recombination and transcriptional alterations (Marinus, 1996; 2000). Dam also contributes to the timing at which chromosome replication is initiated (Marinus, 1996). These processes are all regulated by the

Accepted 18 April, 2002. \*For correspondence at the Research and Education Center for Genetic Information. E-mail hmori@gic.aist-nara.ac.jp; Tel. (+81) 743 72 5660; Fax (+81) 743 72 5669.

hemimethylation of the DNA that occurs following the synthesis of a new DNA strand after the passage of the replication fork.

The protection analysis of the whole *E. coli* genome has revealed that global regulators, including CRP, Fnr and IHF, can block Dam-mediated methylation of many GATC sequences (Wang and Church, 1992; Tavazoie and Church, 1998). This suggests that the recognition sites of these global regulators may coincide with GATC sequences, and that Dam methylation may serve to limit the access of global regulators to upstream regions of a gene, thereby regulating transcription. Observations with individual genes support this notion. For example, the DNA methylation pattern of the two GATC sites within the regulatory regions of the pyelonephritis-associated pilus (*pap*) operon controls *pap* transcription because it affects the ability of two regulatory proteins [leucine-responsive regulatory protein (Lrp) and *pap*-encoded co-regulatory protein (PapI)] to bind upstream (Blyn *et al.*, 1990; Braaten *et al.*, 1994; Nou *et al.*, 1995; van der Woude *et al.*, 1998). Furthermore, computational analysis of part of the *E. coli* genome has revealed that GATC sequences have an unusual distribution in that they often cluster within Fnr and CRP recognition sequences located upstream of respiratory or DNA replication genes (Henaut *et al.*, 1996). Thus, Dam-mediated GATC methylation may affect protein–DNA interaction by modifying the recognition sequence of transcriptional regulators or RNA polymerases (Marinus, 1996). This suggests that many genes with GATC sequences in or near recognition sites for transcriptional regulators may be regulated by Dam-mediated GATC methylation. However, the exact mechanism by which Dam regulates transcription and the extent of its biological importance remains unclear.

It has been reported recently that Dam participates in the virulence of *Salmonella typhimurium* (Heithoff *et al.*, 1999), as Dam-deficient *S. typhimurium* mutants can colonize mucosal sites but are unable to penetrate deeper into the tissue. The Dam-deficient mutants also cannot invade non-phagocytic cells, a function that is required for the virulence of *S. typhimurium*, although normal intracellular proliferation was observed. It was found that Dam controls the expression of a large number of genes that may participate in the invasion of host cells. However, how Dam does this and which genes are particularly crucial for the invasion of host cells by *S. typhimurium* is still not clear.

In this study, we performed a computational analysis of the 500 bp upstream of the ORFs of all *E. coli* genes. We found that, as described previously by Henaut *et al.* (1996), GATC sequences/sites are not randomly distributed, and often overlap with sequences recognized by global regulators. This supports the notion that Dam is important in global gene regulation. We then assessed *E.*

*coli* gene expression of a *dam* mutant under various environmental conditions using transcriptome and proteome techniques. This showed that Dam up- and downregulates many genes, including genes involved in energy and nucleotide metabolism and cellular processes, as well as SOS and stress response genes and translation-related genes. In addition, as genes involved in *E. coli* motility also appear to be controlled by Dam, we speculate that the poor virulence of the Dam mutant *S. typhimurium* may result from defects in bacterial motility.

## Results and discussion

### *GATC sites occur frequently in the transcriptional regulatory region of genes in the entire E. coli genome*

It has been suggested that GATC sites are part of the sequences recognized by the global regulators CRP, Fnr and IHF (Wang and Church, 1992; Hale *et al.*, 1994; Henaut *et al.*, 1996; Tavazoie and Church, 1998). To assess the GATC distribution in the *E. coli* genome comprehensively, we performed computational analysis on the whole *E. coli* genome. This was done by examining the 500 bp upstream of all the ORF start points in the entire *E. coli* genome for GATC sites. If the A, T, G and C residues in *E. coli* genomic sequences were to be purely randomly distributed, the average distance between separate GATC sequences would be about 256 bp (i.e. less than two GATC sequences should be present in each 500 bp upstream sequence; Henaut *et al.*, 1996). However, we found that about 50% of all *E. coli* genes contained more than two GATCs in the 500 bp upstream of the ORF start, with 630 genes and 222 genes containing three and more than five GATCs respectively. The genes with the highest number of GATC sites were *yahG* and *gidA*, which contain 10 and 18 GATCs respectively. There were also regions of low GATC density, the *rhs* gene family being located in such a region. The *rhs* family is large and consists of imperfectly repeated DNA sequences. This family is responsible for duplication within the *E. coli* chromosome (Lin *et al.*, 1984). The results of our computational sequence analysis of the entire *E. coli* genome are shown on our web site ([http://ecoli.aist-nara.ac.jp/xp\\_analysis/dam/all.html](http://ecoli.aist-nara.ac.jp/xp_analysis/dam/all.html)). This non-random and frequent localization of GATC in the promoter regions of the *E. coli* genome suggests that this sequence may participate in gene expression and/or genome structure.

We also confirmed that GATC sites overlap with the consensus sequences for the global regulators Fnr and/or CRP that are found in the regulatory region of several genes/operons (data not shown; Henaut *et al.*, 1996). This supports the hypothesis that GATC may be part of the *cis*-acting elements that are bound by these global transcriptional regulators, and that its methylation affects the DNA

recognition by these regulators. To explore further the biological role of Dam-mediated GATC methylation, we used DNA microarray and two-dimensional PAGE techniques to compare the expression of the whole genome of wild-type *E. coli* and an isogenic Dam-deficient mutant that were grown under various growth conditions.

#### *Microarray analysis of gene expression in the dam-16::kam mutant*

The *dam* mutant, which is defective in Dam methylase activity, is *dam-16::kam*, the result of inserting the kanamycin cassette in the opposite direction to *dam* gene transcription (Parker and Marinus, 1988). The wild-type and mutant cells were grown in L broth under aerobic or low aerobic conditions and sampled in both logarithmic (log) and stationary phases. Total RNA was prepared twice from independently grown cultures, and each preparation was used twice for the hybridization analysis. Thus, for each gene in each strain, transcription levels are represented by four independent measurements (see *Experimental procedures*). The transcription levels of the genes in the *dam* mutant are expressed relative to those of the wild type, yielding a relative ratio (see *Experimental procedures*). The genes with significantly altered transcriptional levels in the *dam* mutant cells are summarized in Table 1. Supplemental data are available on our web site ([http://ecoli.aist-nara.ac.jp/xp\\_analysis/dam/all.html](http://ecoli.aist-nara.ac.jp/xp_analysis/dam/all.html)).

#### *The dam mutation affects the transcriptional expression of energy metabolism and respiratory enzymes*

The genes with altered transcriptional levels included most genes involved in the TCA cycle. Under low aerobic conditions, *acnA* (aconitate hydratase; underlined genes in the text are listed in Table 1), *gltA* (citrate synthase) and the *sucACD* (2-ketoglutarate dehydrogenase, succinyl-CoA synthetase( $\alpha,\beta$ )) operon were transcribed at lower levels in the *dam* mutant than in the wild type (Table 1). Furthermore, under aerobic conditions, the expression levels of the *aceBA* operon (whose operon is controlled by global regulators IHF, IclR and FadR; Resnik *et al.*, 1996) and *fumA* were increased in the *dam* mutant. In addition, several TCA-related genes (e.g. *sdhCDB*, *sucAB* and *lpdA*) were also moderately upregulated in the *dam* mutant under aerobic conditions (data not shown, see our web site [http://ecoli.aist-nara.ac.jp/xp\\_analysis/dam/all.html](http://ecoli.aist-nara.ac.jp/xp_analysis/dam/all.html)). Thus, in normal *E. coli* grown in aerobic conditions, the cellular levels of TCA cycle enzymes are directly or indirectly negatively controlled at the transcriptional level by GATC methylation, whereas in low aerobic conditions, the expression is positively regulated.

Also affected by the *dam* mutation was the regulation of genes involved in sugar metabolism and degradation (*araC* and the *srIEABD-gutM-srIRQ* operon, whose tran-

scription is controlled by CRP and *ppsA*), the degradation of phospholipid (*fadB* and *ycfX*, whose transcription is controlled by CRP, see Fig. 4) and the metabolism of carbohydrates (*mhpE* and *mhpR*) as, under aerobic conditions, the transcription of these genes was increased in the *dam* mutant. In contrast, in *dam* mutant cells grown to log phase under aerobic conditions, the transcription of several respiratory enzyme genes, especially related to anaerobic respiration, was decreased. These genes included *dmsA* and *frdB* (whose synthesis is controlled by Fnr), the *nirBDC* operon (encoding anaerobic reductase), the *nap* operon *napFDAGHBC* (encoding periplasmic nitrate reductase), the *narGHJ* operon (encoding the major membrane-bound nitrate reductase) and the *narK* and *Z* genes, which encode the nitrate transport protein and the alpha-chain of respiratory nitrate reductase 2 respectively. The *nar* operon is controlled by Fnr or NarL (Li *et al.*, 1994). Furthermore, several genes of anaerobic respiration-related genes were also moderately downregulated in a *dam* mutant under low aerobic conditions (e.g. *narG*, *narJ* and *narZ*; data not shown). It suggested that the expression of anaerobic respiration-related genes were decreased in the *dam* mutant.

Several of the genes were upregulated by transcription in the *dam* mutant grown in both aerobic and low aerobic conditions. These genes include *yahK* (hypothetical zinc-type alcohol dehydrogenase), *citE* (citrate lyase alpha chain), *mhpE* (4-hydroxy-2-oxovalerate aldolase) and *ppsA* (phosphoenolpyruvate synthase, whose expression is positively regulated by FruR). In contrast, transcription of the *gatABCDXYZ* operon, whose products convert galactitol to dihydroxyacetone-P and glyceraldehyde-3-P (PTS system), was downregulated in the mutant strain in aerobic and low aerobic conditions.

These observations together suggest that Dam-dependent transcriptional control may participate in energy metabolism and respiration by regulating the control of global regulators.

#### *The dam mutation affects the transcription of genes involved in cofactor biosynthesis and metabolism of amino acids, fatty acids phospholipids and nucleotides*

In log phase *dam* mutant bacteria grown in aerobic conditions, the transcription of several genes involved in amino acid metabolism was elevated. These genes include *avtA*, the *dadA,X* cluster, the *hisGDC* operon, *ltaA*, *putA* and the *thrABC* operon. The transcription of several other genes was also elevated in these conditions. These included *entE* involved in enterobactin synthesis, the genes encoding pantothenate synthase (*panD* and *C* cluster) and genes participating in riboflavin synthesis (*ribD* and *H* cluster). Transcription of genes involved in lipid (*accC*, *fabB* and *I*) and phosphatidic acid

**Table 1.** Genes differentially expressed between KK46(*dam+*) and KK335(*dam-16::kam*).

	Relative log ratio (KK335( <i>dam-16::kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
<b>Amino acid metabolism</b>						
<i>ansB</i>	<b>-0.88</b>				Asparaginase (EC 3.5.1.1) II precursor	2
<i>argM</i>			-1.19		Succinylornithine aminotransferase, N-(alpha)-acetylornithine-(delta)-aminotransferase (EC 2.6.1.-) (Carbon starvation protein C)	2
<i>astA</i>			-1.37		Arginine succinyltransferase	3
<i>avtA</i>	<b>0.88</b>		<b>0.71</b>		Valine-pyruvate aminotransferase (EC 2.6.1.66) (Transaminase C) (alanine-valine transaminase)	1
<i>dadA</i>	<b>1.79</b>		<b>-0.85</b>		D-Amino acid dehydrogenase small subunit (EC 1.4.99.1)	1
<i>dadX</i>	<b>1.55</b>				Alanine racemase, catabolic precursor (EC 5.1.1.1)	4
<i>dapB</i>	<b>1.36</b>	<b>-0.74</b>			Dihydrodipicolinate reductase (EC 1.3.1.26)	2
<i>dpaL</i>		1.53			Putative diamino propionate ammonia lyase[EC4.3.1.15] [diaminopropio-natase] (Alpha, beta-diaminopropionate ammonia-lyase)	3
<i>gltD</i>			<b>-0.83</b>	<b>0.72</b>	Glutamate synthase (NADPH) small chain (EC 1.4.1.13) (glutamate synthase beta subunit) (nadph-gogat) (GltS beta chain)	1
( <i>gltD</i> )		1.26			Hypothetical protein	1
<i>gph</i>	<b>-1.01</b>	<b>-0.45</b>	<b>-1.59</b>		Phosphoglycolate phosphatase (EC 3.1.3.18) (PGP)	2
<i>hisC</i>	1.07				Histidinol-phosphate aminotransferase	3
<i>hisD</i>	<b>0.74</b>				Histidinol dehydrogenase (EC 1.1.1.23) (hdh)	4
<i>hisG</i>	<b>0.73</b>				ATP phosphoribosyltransferase	3
<i>ltaA</i>	<b>2.53</b>	<b>1.06</b>	<b>2.16</b>		L-Allo-threonine aldolase (EC 4.1.2.-)	2
<i>putA</i>	<b>0.72</b>				Proline dehydrogenase (EC 1.5.99.8) (proline oxidase)/delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12) (p5c dehydrogenase)	0
<i>sdaB</i>	<b>-0.76</b>				L-Serine deaminase	1
<i>sol</i>	<b>0.99</b>				Sarcosine oxidase (EC 1.5.3.1)	0
<i>speA</i>	<b>0.74</b>		<b>0.69</b>		Arginine decarboxylase (EC 4.1.1.19)	0
<i>tdcB</i>			-1.8		Threonine dehydratase catabolic (EC 4.2.1.16) (threonine deaminase)	0
<i>thrA</i>	<b>1.09</b>				ThrA bifunctional enzyme, aspartokinase I bifunctional enzyme N-terminal is aspartokinase I and C-terminal is homoserine dehydrogenase I (EC 2.7.2.4) (EC 1.1.1.3)	0
<i>thrC</i>	<b>0.97</b>		<b>-1.58</b>		Threonine synthase (EC 4.2.99.2)	1
<b>Biosynthesis of cofactors, prosthetic groups, carriers</b>						
<i>bioA</i>	>1.91		>3.33		Adenosylmethionine-8-amino-7-oxononanoate transaminase (EC 2.6.1.62)	3
( <i>bioD</i> )	<b>-1.37</b>				Dethiobiotin synthase (EC 6.3.3.3)	1
<i>entE</i>	1.86		>1.75		Enterochelin synthetase (EC 6.-.-.-) component E	2
<i>ispA</i>	<b>1.28</b>				Geranyl transtransferase (EC 2.5.1.10) (farnesyl-diphosphate synthase)(FPP synthase)	0
<i>nadD</i>				-0.97	Nicotinic acid mononucleotide adenyllyltransferase, NAMN adenyllyltransferase	0
<i>panC</i>	0.94				Pantoate-beta-alanine ligase (EC 6.3.2.1) (pantothenate synthetase) (pantoate activating enzyme)	3
<i>panD</i>	1.08				Aspartate 1-decarboxylase (EC 4.1.1.11)	0
<i>ribD</i>	<b>1.17</b>	<b>0.86</b>	<b>0.89</b>	<b>0.94</b>	Riboflavin-specific deaminase	1
<i>ribH</i>	<b>0.74</b>				6,7-Dimethyl-8-ribityllumazine synthase	2
<i>yncB</i>				-1.63	Putative NADP-dependent oxidoreductase (EC 1.-.-.-)	1
<b>Cell envelope</b>						
<i>crl</i>	<b>0.83</b>			<b>-0.86</b>	Crl protein	3
<i>cheA</i>			-0.93		Chemotaxis protein CheA (EC 2.7.3.-), CheY kinase	3
<i>cheR</i>			-1.16		Chemotaxis protein methyltransferase (EC 2.1.1.80)	4
<i>cheY</i>			-1.26		Chemotaxis protein CheY, chemotaxis response regulator protein	4
<i>cheZ</i>			-1.16		CheY phosphatase	1
<i>fimB</i>	<b>-1.05</b>			<b>0.97</b>	Type 1 fimbriae regulatory protein FimB	0
<i>flgD</i>		<b>0.78</b>			Basal-body rod modification protein	0
<i>fliC</i>			-1.4		Flagellin	1
<i>fliK</i>	<b>0.84</b>				Flagellar hook-length control protein	3
<i>motB</i>			-1.05		Chemotaxis MotB protein (motility protein b)	1

Table 1. cont.

	Relative log ratio (KK335( <i>dam-16::kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
<i>ompA</i>				<b>-0.99</b>	Outer membrane protein a precursor (outer membrane protein II)	1
<i>ompX</i>	<b>-0.69</b>			<b>-0.89</b>	Outer membrane protease	0
<i>ompW</i>	-1.66				Outer membrane protein W	3
<i>rfbX</i>	<b>-0.85</b>				O-antigen transporter	0
<i>rfc</i>	<b>-1.48</b>	<b>-0.89</b>	<b>-0.93</b>		Probable O-antigen polymerase	1
<i>slp</i>				1.2	Slp protein	0
<i>tap</i>		-1.11	-1.63		Methyl-accepting chemotaxis protein II (Mcp-II) (aspartate chemoreceptor protein)	4
<i>tar</i>			-1.45		Methyl-accepting chemotaxis protein II (Mcp-II) (aspartate chemoreceptor protein)	2
<i>trg</i>				-0.99	Methyl-accepting chemotaxis protein III (Mcp-III) (ribose and galactose chemoreceptor protein)	1
Central intermediary metabolism						
<i>gabD</i>				-1.26	Succinate-semi-aldehyde dehydrogenase (NADP <sup>+</sup> ) (EC 1.2.1.16) (ssdh)	2
<i>gabT</i>				-0.88	4-Aminobutyrate transaminase (EC 2.6.1.19)	1
<i>glpK</i>	<b>0.76</b>				Glycerol kinase (EC 2.7.1.30)	2
<i>hdhA</i>	<b>-0.78</b>			<b>-0.75</b>	7-Alpha-hydroxysteroid dehydrogenase (EC 1.1.1.159) (7-alpha-hsdH)	2
<i>nagA</i>			-1.19		N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) (NagA)	1
<i>nagB</i>			-1.37		Glucosamine-6-phosphate isomerase (EC 5.3.1.10) (glucosamine-6-phosphate deaminase)	1
<i>nagC</i>			-0.86		N-acetylglucosamine repressor	1
<i>nanA</i>	<b>-1.92</b>		<b>-2.7</b>		N-acetylneuraminase lyase subunit (EC 4.1.3.3) (N-acetylneuraminic acid aldolase) (N-acetylneuraminase pyruvate lyase) (nalase)	9
<i>nanE</i>			-2.04		ManNAc epimerase	5
<i>nanK</i>	<b>-1.08</b>		<b>-1.21</b>		ManNAc kinase	2
<i>rfbC</i>		<b>-0.76</b>			dTDP-6-deoxy-D-glucose-3,5 epimerase	1
<i>yojH</i>	1.2				Malate:quinone oxidoreductase	3
Energy metabolism						
<i>aceA</i>	<b>0.78</b>				Isocitrate lyase (EC 4.1.3.1) (isocitrase) (isocitratase) (icl)	5
<i>aceB</i>	<b>0.85</b>				Malate synthase a (EC 4.1.3.2) (msa)	6
<i>ackA</i>			<b>0.79</b>		Acetate kinase (EC 2.7.2.1)	2
<i>acnA</i>				-1.2	Aconitate hydratase (EC 4.2.1.3)	4
<i>adhE</i>			-0.99		Alcohol dehydrogenase (EC 1.1.1.1)	1
<i>araC</i>	1.39				Arabinose operon regulatory protein	2
<i>astB</i>				-1.11	Succinylarginine dihydrolase	2
<i>astD</i>				-1.17	Succinylglutamic semi-aldehyde dehydrogenase	2
<i>citE</i>	1.32		1.44		Citrate lyase beta chain (acyl lyase subunit) (citE) homologue	1
<i>cydB</i>				<b>-0.99</b>	Cytochrome <i>d</i> ubiquinol oxidase subunit II (EC 1.10.3.-)	1
<i>cyoA</i>				-1.09	Cytochrome O ubiquinol oxidase (EC 1.10.-.-) chain II	1
<i>cyoD</i>		<b>0.79</b>			CyoD protein	2
<i>dmsA</i>	<b>-1.43</b>				Dimethyl sulphoxide reductase chain A	1
<i>fadB</i>	<b>1.24</b>				Fatty oxidation complex alpha subunit (contain: enoyl-CoA hydratase (EC 4.2.1.17), delta(3)- <i>cis</i> -delta(2)- <i>trans</i> -enoyl-CoA isomerase (EC 5.3.3.8), 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35), and 3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3))	4
<i>frdB</i>	<b>-0.77</b>				Fumarate reductase (EC 1.3.99.1) iron-sulphur protein	2
<i>fucR</i>			-1.43		Fuc operon regulatory protein	3
<i>fumA</i>	<b>1.13</b>				Fumarate hydratase (EC 4.2.1.2) FumC, iron dependent	1
<i>gatA</i>			-1.17		PTS system, galactitol-specific IIA component	1
<i>gatB</i>			<b>-2.18</b>		Phosphotransferase system enzyme II, galactitol specific, protein B	1
<i>gatC</i>			-2.16		PTS system, galactitol-specific IIC component (EIIC-GAT) (galactitol-permease IIC component) (phosphotransferase enzyme II, C component)	1
<i>gatD</i>	-1.43		<-3.51		Galactitol-1-phosphate 5-dehydrogenase (EC 1.1.1.251)	3

Table 1. cont.

	Relative log ratio (KK335( <i>dam-16::kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
<i>(gatR)</i>			-1.24		Galactitol utilization operon repressor	2
<i>gatY</i>	<b>-0.94</b>		<b>-1.52</b>	<b>-0.81</b>	Tagatose-bisphosphate aldolase (EC 4.1.2.-)	1
<i>gatZ</i>			-1.56		Putative tagatose 6-phosphate kinase (EC 2.7.1.143)	1
<i>glf</i>	<b>-0.96</b>	<b>-0.66</b>			UDP-galactopyranose mutase (EC 5.4.99.9)	0
<i>glgS</i>				-1.22	RpoS-dependent glycogen synthesis protein	2
<i>glk</i>			-1.11		Glucokinase (EC 2.7.1.2)	3
<i>gltA</i>			-0.92		Citrate synthase (EC 4.1.3.7)	2
<i>gutD</i>	1.03				Sorbitol-6-phosphate 2-dehydrogenase (EC 1.1.1.140) (glucitol-6-phosphate dehydrogenase) (ketosephosphate reductase)	2
<i>gutM</i>	1.35				Glucitol operon activator protein	7
<i>kdsA</i>	1.16				3-deoxy-D-manno-octulosonic acid 8-phosphate synthetase	5
<i>hoxK</i>	-1.11				Hydrogenase (EC 1.18.99.1) small-chain precursor	3
<i>lldD</i>			-1.04		L-Lactate dehydrogenase (cytochrome) (EC 1.1.2.3)	2
<i>mhpE</i>	1.94		1.84		4-Hydroxy-2-oxovalerate aldolase (EC 4.1.3.-)	0
<i>mhpR</i>	2.86				Mhp operon transcriptional activator	2
<i>napA</i>	-1.66				Probable periplasmic nitrate reductase 3 (EC 1.7.99.4)	2
<i>napD</i>	-1.24				NapD protein	1
<i>napF</i>	-1.54				Ferredoxin-type protein NapF	2
<i>narG</i>	-2.25	-0.88			Respiratory nitrate reductase 1 alpha chain (EC 1.7.99.4)	1
<i>narH</i>		-0.99			Respiratory nitrate reductase 1 beta chain (EC 1.7.99.4)	2
<i>narJ</i>			<-1.02		NarJ protein	1
<i>narK</i>	<b>-1.31</b>				Nitrate transport protein NarK	2
<i>narZ</i>	-1.77	-0.94			Respiratory nitrate reductase 2 alpha chain (EC 1.7.99.4)	3
<i>nirB</i>	<b>-2.19</b>		<b>-1.51</b>		Nitrite reductase (NAD(P)H) large subunit (EC 1.6.6.4)	1
<i>nirD</i>			-1.56		Nitrite reductase (NAD(P)H) small subunit (EC 1.6.6.4)	4
<i>nuoH</i>			-0.8		NADH dehydrogenase I chain H (EC 1.6.5.3) (NADH-ubiquinone oxidoreductase chain 8) (nuo8)	2
<i>ppsA</i>	<b>1.52</b>		<b>1.32</b>		Pyruvate, water dikinase (EC 2.7.9.2)	3
<i>rpe</i>	-1.12		-1.22		Ribulose-phosphate 3-epimerase (EC 5.1.3.1) (pentose-5-phosphate 3-epimerase) (ppe)	3
<i>sfsA</i>				-1.04	Sugar fermentation stimulation protein	0
<i>sucA</i>			-1.06		Oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2)	1
<i>sucC</i>			-1.12		Succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) beta chain	7
<i>sucD</i>			-0.81		Succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) alpha chain	3
<i>tpiA</i>	<b>0.72</b>				Triosephosphate isomerase (EC 5.3.1.1) (tim)	1
<i>treC</i>	<b>-0.74</b>		<b>-1.68</b>		Trehalose-6-phosphate hydrolase (EC 3.2.1.93) (alpha, alpha-phosphotrehalase)	1
<i>udhA</i>			-1.37		Soluble pyridine nucleotide transhydrogenase	1
<i>yahK</i>	G2(up)		>1.32		Hypothetical zinc-type alcohol dehydrogenase-like protein	2
<i>(yagR)</i>		1.62			Hypothetical oxidoreductase protein	2
<i>ydbC</i>			-0.94		Hypothetical oxidoreductase	1
<i>ydeV</i>			-0.83		Hypothetical sugar kinase	1
<i>yfcX</i>	<b>1.84</b>	<b>1.13</b>			Putative fatty oxidation complex alpha subunit (enoyl-CoA: hydratase (EC 4.2.1.17))	3
<i>yibO</i>	<b>-0.81</b>		<b>-0.84</b>		Putative 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1) (phosphoglyceromutase)	3
<i>yjiY</i>	-1.01		-1.54		Carbon starvation protein A homologue	0
Environmental, metabolic response						
<i>bolA</i>			-1.29		BolA protein	4
<i>cspD</i>			-0.92		Cold shock-like protein CspD	1
<i>cspE</i>	1.3				CspE protein	0
<i>dnaJ</i>	1.01				DnaJ protein, heat shock protein	0
<i>dnaK</i>	<b>0.76</b>		<b>-0.87</b>		DnaK protein, chaperone Hsp70	1
<i>dppA</i>	<b>1.44</b>				Dipeptide-binding protein DppA precursor	1
<i>hslU</i>	<b>1.17</b>				Heat shock protein HslU	3
<i>hslV</i>	<b>1.49</b>				Heat shock protein HslV	1
<i>htpG</i>			<b>-0.78</b>		Heat shock protein C62.5	3
<i>ibpA</i>	<b>1.66</b>				16 kDa heat shock protein A	2
<i>lytB</i>	<b>0.73</b>				Penicillin tolerance protein (lytB), probable metalloproteinase	2
<i>groES</i>	<b>0.94</b>				10 kDa chaperonin (protein CPN10) (protein GroES)	1
<i>osmB</i>			-1.2		Lipoprotein OsmB precursor, osmotically inducible	1

Table 1. cont.

	Relative log ratio (KK335( <i>dam-16::kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
<i>rnb</i>	1.21				Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II)	1
<i>sfmC</i>	<b>2.54</b>	<b>0.77</b>	<b>2.42</b>	<b>0.77</b>	Chaperone protein SmfC precursor	1
<i>sulA</i>	<b>1.57</b>	<b>2.07</b>	<b>1.28</b>	<b>1.21</b>	Cell division inhibitor	5
<i>surA</i>	<b>0.97</b>				Survival protein SurA precursor (peptidyl-prolyl <i>cis-trans</i> isomerase SurA) (EC 5.2.1.8) (PPiase) (rotamase C)	1
<i>yhbU</i>	<-1.20				Putative protease (o331)	1
<i>uspB</i>				-1.06	Universal stress protein B	2
<i>ybeW</i>			-0.77		Putative chaperone protein HscC	4
Fatty acid and phospholipid metabolism						
<i>accC</i>	<b>1.04</b>				Acetyl-CoA carboxylase (EC 6.4.1.2), biotin carboxylase	2
<i>acs</i>				-1.43	Acetyl-coenzyme A synthetase (EC 6.2.1.1) (acetate-CoA ligase) (acyl-activating enzyme)	1
<i>envM</i>	1.44				Enoyl-[acyl-carrier-protein] reductase (NADH) (EC 1.3.1.9) (NADH-dependent enoyl-acyl reductase)	0
<i>fabB</i>	<b>1.04</b>				3-oxoacyl-[acyl-carrier-protein] synthase I (EC 2.3.1.41) (beta-ketoacyl-acyl synthase I) (KasI)	3
<i>fabI</i>	<b>1.42</b>				Enoyl-[acyl-carrier-protein] reductase (NADH) (EC 1.3.1.9) (NADH-dependent enoyl-acyl reductase)	0
<i>flxA</i>			-1.17		Gene whose expression is dependent on the flagellum-specific sigma factor, FlIA, but dispensable for motility development	1
<i>gpsA</i>	<b>0.73</b>				L-Glycerol 3-phosphate dehydrogenase	2
<i>ybbO</i>	<b>0.79</b>				Hypothetical oxidoreductase (EC 1.-.-.)	0
<i>ybhO</i>	<b>3.48</b>	<b>0.98</b>			Hypothetical protein	2
<i>yebF</i>	<b>0.93</b>	<b>1.48</b>	<b>2.23</b>		Hypothetical lipoprotein (ORF122)	1
<i>yefI</i>		<b>-0.76</b>			Hypothetical protein	2
<i>yfcY</i>	1	<b>1.01</b>			Probable 3-ketoacyl-CoA thiolase (EC 2.3.1.16)	0
Nucleotide metabolism						
<i>add</i>	-1.41				Adenosine deaminase (EC 3.5.4.4)	2
<i>adk</i>	<b>1.38</b>				Adenylate kinase (EC 2.7.4.3) (ATP-amp transphosphorylase)	4
<i>apt</i>	1.1		<b>1.15</b>		Adenine phosphoribosyltransferase (EC 2.4.2.7) (AprT)	0
<i>codA</i>	<b>1.24</b>		<b>0.83</b>		Cytosine deaminase (EC 3.5.4.1)	2
<i>gpt</i>	<b>0.92</b>			<b>0.08</b>	Xanthine-guanine phosphoribosyltransferase (EC 2.4.2.22) (XgpRT)	3
<i>gsk</i>	1.19				Inosine-guanosine kinase (EC 2.7.1.73)	2
<i>guaC</i>	<b>1.68</b>				GMP reductase (EC 1.6.6.8)	3
<i>ndk</i>	<b>1.65</b>				Nucleoside-diphosphate kinase (EC 2.7.4.6)	0
<i>nrdD</i>	-1.71				Oxygen-sensitive ribonucleoside-triphosphate reductase (EC 1.17.4.-)	1
<i>prsA</i>	1.21		1.13		Ribose-phosphate pyrophosphokinase (EC 2.7.6.1) (phosphoribosyl pyrophosphate synthetase)	2
<i>purB</i>	<b>1.08</b>		<b>0.96</b>		Adenylosuccinate lyase (EC 4.3.2.2)	2
<i>purD</i>			1.5		Phosphoribosylamine-glycine ligase (EC 6.3.4.13)	1
<i>purE</i>	2.09		2.04		Phosphoribosylaminoimidazole carboxylase (EC 4.1.1.21) catalytic chain	2
<i>purF</i>			1.81		Amidophosphoribosyltransferase (EC 2.4.2.14) (glutamine phosphoribosylpyrophosphate amidotransferase) (atase)	4
<i>purH</i>			1.4		PurH bifunctional enzyme.	1
<i>purK</i>			1.58		Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21) (air carboxylase) (AirC)	0
<i>purT</i>			1.51		Glycinamide ribonucleotide transformylase	1
<i>pyrC</i>	1.15				Dihydroorotase (EC 3.5.2.3)	1
<i>upp</i>	<b>-0.8</b>				Uracil phosphoribosyltransferase (EC 2.4.2.9) (UMP pyrophosphorylase) (uptase)	0
Regulatory functions						
<i>304#1</i>				-0.9	Hypothetical transcriptional regulator	1
<i>cpxP</i>	<b>-0.91</b>			<b>-0.78</b>	Periplasmic protein precursor	1
<i>fnr</i>	<b>0.93</b>				Fumarate and nitrate reduction regulatory protein.	1
<i>himA</i>				-1.22	Integration host factor alpha-subunit (IHf-alpha)/bending	2
<i>hns</i>				-1.27	DNA-binding protein H-NS/bending	0
<i>lacI</i>	<b>1.49</b>		<b>0.84</b>		Lac repressor	3

Table 1. cont.

		Relative log ratio (KK335( <i>dam</i> -16:: <i>kam</i> )/KK46) <sup>a</sup>				
		Aerobic		Low aerobic		
	Log	Stationary	Log	Stationary	Description <sup>b</sup>	No. of GATC <sup>c</sup>
			1.09		LexA repressor (EC 3.4.21.88)	2
				-1.41	RelB protein	1
				-1.24	Hypothetical protein	1
	>2.03		2.65		Hypothetical transcriptional regulator	6
	>3.03	2.16	2.81		Hypothetical transcriptional regulator	2
		-0.83			Hypothetical transcriptional regulator	2
			0.94		Regulator protein of zinc uptake system znuABC	0
Replication						
			1.51		DNA damage-inducible protein F	1
	<b>1.38</b>	<b>1.48</b>	<b>1.6</b>	<b>0.77</b>	DNA damage-inducible protein I	0
				-1.01	Dosage-dependent DnaK suppressor protein	1
		0.96			DNA gyrase inhibitory protein	0
	1.2				DNA-directed DNA polymerase (EC 2.7.7.7) III delta chain	5
	<b>1.19</b>				Transcription-repair coupling protein Mfd	2
	<b>1.21</b>				Primosomal protein 'n' precursor	0
		1.45	1.15	1.04	ATP-dependent recombinase	1
	<b>1.8</b>	<b>2.01</b>	<b>1.97</b>	<b>1.43</b>	DNA repair protein RecN (recombination protein n)	3
		<b>1.03</b>			Recombination protein RecR	3
	<b>0.77</b>		<b>0.88</b>		RuvA protein	1
	1.07				Excinuclease ABC subunit A	1
Transcription						
	<b>-0.73</b>		<b>0.96</b>		ATP-dependent RNA helicase	1
	<b>-0.98</b>				N utilization substance protein A (NusA protein) (I factor)	3
	<b>0.91</b>			<b>-0.94</b>	N utilization substance protein B (NusB protein)	1
	1.85		1.74		Putative ATP-dependent RNA helicase RhIE	0
	<b>-0.96</b>		<b>-1.15</b>	<b>-1.2</b>	Hypothetical protein (URF1) (ORFS54)	2
Translation						
	1.03				Asparaginyl-tRNA synthetase (EC 6.1.1.22) (asparagine-tRNA ligase) (asnRS)	1
		<b>0.78</b>	<b>0.81</b>		Glutamyl-tRNA synthetase (EC 6.1.1.18) (glutamine-tRNA ligase) (GlnRS)	5
				-1.56	Osmolarity effect protein	2
				<b>-0.78</b>	Translation initiation factor IF-1	1
				<b>-0.73</b>	Initiation factor IF-3	1
	-1.09				Ribosome-binding factor a (p15b protein)	3
	<b>-1.06</b>				Ribosomal protein L1	0
	<b>-0.74</b>				50S ribosomal protein L5	0
	<b>-0.73</b>				Ribosomal protein L11	2
			0.98		Ribosomal protein L13	2
	<b>-0.73</b>				Ribosomal protein I19	1
	<b>-0.71</b>				50S ribosomal protein L23	5
	<b>-0.94</b>				Ribosomal protein S15.	3
	<b>-0.82</b>				Ribosomal protein S19	1
			1.18		rRNA(guanine-N1-)-methyltransferase	2
Transport and binding protein						
	<b>-1.02</b>		<b>-1.7</b>		Putative tagatose 6-phosphate kinase AgaZ (EC 2.7.1.-)	7
				-1.17	Lysine-arginine-ornithine-binding periplasmic protein precursor (lao-binding protein)	2
			<b>0.95</b>		DctA protein	1
	<b>0.82</b>				Transmembrane protein DppB	3
	<b>0.95</b>	<b>0.7</b>			Dipeptide transport ATP-binding protein DppF	1
	<b>-2.36</b>			<b>-0.72</b>	Probable formate transporter	3
	<b>-0.71</b>				Ferritin	4
	<b>0.83</b>				Glutamine-binding protein precursor	1
				-1.01	Phosphotransferase system enzyme II (EC 2.7.1.69), glucose-specific, factor III	1
				-1.05	Alpha-ketoglutarate permease	0
			-2.73		Maltoporin precursor (lambda receptor protein)	2
			-2.9		Maltose-binding protein precursor	1
			-2.9		Inner membrane protein MalF	2



Table 1. cont.

	Relative log ratio (KK335( <i>dam</i> -16:: <i>kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
<i>malG</i>			-1.83		Maltose transport protein MalG	5
<i>malM</i>			-2.12		Maltose operon periplasmic protein precursor	0
<i>malK</i>			-2.59		Maltose/maltodextrin transport ATP-binding protein MalK	1
<i>malP</i>			-1.74		Maltodextrin phosphorylase (EC 2.4.1.1)	3
<i>malS</i>			<-2.53		Alpha-amylase (EC 3.2.1.1) precursor	1
<i>mgIB</i>			-2.12		D-galactose-binding protein precursor	2
<i>malQ</i>			-1.56		4-Alpha-glucanotransferase (EC 2.4.1.25) (amylomaltase) (disproportionating enzyme) (d-enzyme)	3
<i>manX</i>	<b>-0.83</b>	<b>-0.8</b>	<b>-1.46</b>	<b>-1.04</b>	Phosphotransferase system enzyme II (EC 2.7.1.69), mannose-specific, factor III	2
<i>manY</i>			-1.39		Phosphotransferase system enzyme II (EC 2.7.1.69), mannose-specific, factor II-P	3
<i>manZ</i>			-1.56	-0.99	PTS system, mannose-specific IID component (EIID-Man) (mannose-permease IID component) (phosphotransferase enzyme II, D component) (EII-M-Man)	3
<i>modA</i>	<b>-0.77</b>		<b>-0.85</b>		Molybdate-binding periplasmic protein precursor	0
<i>nagE</i>			-0.98		PTS system, N-acetylglucosamine-specific IIABC component (EIIABC-Nag) (N-acetylglucosamine-permease IIABC component) (phosphotransferase enzyme II, ABC component) (EC 2.7.1.69) (EII-Nag)	3
<i>nanT</i>	<b>-1.77</b>		<b>-2.42</b>		Putative sialic acid transporter	5
<i>nikA</i>	-1.63				NikA protein.	5
<i>nupC</i>	<b>0.83</b>			<b>0.74</b>	Nucleoside permease (nucleoside-transport system protein)	3
( <i>potC</i> )				-1.11	Spermidine/putrescine transmembrane protein C	2
<i>proX</i>		0.93			Glycine betaine-binding periplasmic protein precursor.	4
<i>ptsH</i>	<b>-0.81</b>		<b>-0.99</b>	<b>-0.76</b>	Phosphocarrier protein Hpr (histidine-containing protein).	1
<i>putP</i>	<b>0.94</b>	<b>0.78</b>			Proline carrier protein	0
<i>sapC</i>	1.1				Peptide transport system permease protein SapC	4
<i>sapF</i>	1.27				Peptide transport system ATP-binding protein	1
<i>sprE</i>				-0.92	Putative two-component response regulator	1
<i>srlA</i>	<b>1</b>	<b>0.75</b>			PTS system, glucitol/sorbitol-specific IIBC component (EIIIBC-Gut) (glucitol/sorbitol-permease IIBC component) (phosphotransferase enzyme II, BC component) (EC 2.7.1.69) (EII-Gut)	2
<i>treB</i>	<b>-1.04</b>	<b>-0.87</b>	<b>-1.83</b>		Phosphotransferase system trehalose permease	5
<i>xylF</i>		-0.93			D-Xylose-binding periplasmic protein (precursor)	2
<i>yaaA</i>			-0.83		Putative inner membrane transport protein	4
<i>ybhL</i>	<b>1.41</b>		<b>1.43</b>		Probable transport permease	2
<i>ycjV</i>			-0.9		Hypothetical ABC transporter ATP-binding protein	3
<i>ydeA</i>		1.12			L-Arabinose and IPTG exporter protein	3
<i>ygfO</i>		1.35			Hypothetical purine permease	3
Other categories						
<i>ampG</i>	1.22				AmpG protein, regulates beta-lactamase synthesis	3
<i>cinA</i>	-1.43				Putative competence-damage protein	1
<i>emrA</i>		1.16			Multidrug-resistant protein EmrA	5
<i>fms</i>			-1.04		Polypeptide deformylase (EC 3.5.1.31)	0
<i>intD</i>	-1.17				Prophage dLp12 integrase (prophage <i>qsr'</i> integrase)	0
<i>intE</i>		2.54		>1.62	Prophage lambda integrase	1
<i>pqiA</i>	1.73		1.76		Paraquat-inducible protein A	2
<i>yjiY</i>	<b>-1.01</b>		<b>-1.57</b>		Carbon starvation protein A homologue	0
<i>yjiW</i>	<b>-0.76</b>		<b>-0.88</b>		Hypothetical protein	4
VXISBPP21			>1.48		Excisionase	1
Y genes						
129#5			1.7		Hypothetical protein (o198)	2
334#5.1				-1.35	Hypothetical protein	0
356#8		0.94			Hypothetical protein	1
358#3				-0.93	Hypothetical protein	1
411#1	-1.02				Hypothetical protein	3
421#4	<b>-0.91</b>				Hypothetical protein	7
443#3				-1.74	Hypothetical protein	1
467#1		1.29			Hypothetical protein	1
502#6			-1.12		Hypothetical protein	1

Table 1. cont.

	Relative log ratio (KK335( <i>dam</i> -16:: <i>kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
576#14				-1.33	Hypothetical protein	2
<i>yabP</i>	<-0.97				Hypothetical protein	1
<i>yadF</i>	<b>1.15</b>		<b>0.62</b>		Hypothetical protein	2
<i>yadR</i>				-1.22	Hypothetical protein (ORF118)	1
<i>yaeH</i>				-0.94	Hypothetical protein	3
<i>yahO</i>		-1.04		-1.86	Hypothetical protein	2
<i>ybaJ</i>				-1.58	Hypothetical protein	1
<i>ybaX</i>	1.46				Hypothetical protein	1
<i>ybbN</i>	<b>1.29</b>				Hypothetical protein	1
<i>ybcI</i>	4.54	G2(up)	3.11		Hypothetical protein	2
<i>ybdQ</i>	-1.27			-1.16	Unknown protein from 2-D PAGE (spots pr25/lm16/2d_000lr3)	2
<i>ybeY</i>	<b>0.67</b>			<b>-0.81</b>	Hypothetical protein	2
<i>ybhN</i>	2.58		>1.66		Hypothetical protein	7
<i>ybhO</i>	>3.43		>3.45		Hypothetical protein	2
<i>ybiJ</i>	G2(up)				Hypothetical protein	1
<i>ycdQ</i>	<b>1.74</b>	<b>1.29</b>	<b>2.63</b>		Hypothetical protein	0
<i>yceD</i>	<b>-0.85</b>				Hypothetical protein	1
<i>yceP</i>				-1.14	Hypothetical protein	0
<i>ycgB</i>				-1.04	Hypothetical protein	1
<i>ychF</i>	<b>0.8</b>		<b>0.7</b>		Probable GTP-binding protein (ORF-3)	1
<i>ychH</i>	-1.54			-1.51	Hypothetical protein (ORF-2)	1
<i>ychM</i>			<b>0.71</b>		Hypothetical protein	3
<i>yciM</i>	0.95				Hypothetical protein	2
<i>ycjX</i>	1.85				Hypothetical protein	1
<i>ydaA</i>				-0.93	Hypothetical protein	3
<i>ydgA</i>				<b>-0.92</b>	Hypothetical protein (o490)	0
<i>ydhV</i>			-1.77		Hypothetical protein	2
<i>ydiJ</i>	<b>0.9</b>				Hypothetical protein	0
<i>yeaA</i>				<b>-1.06</b>	Hypothetical protein	2
<i>yeaG</i>				-1.09	Hypothetical protein kinase	0
<i>yeaH</i>				-1.02	Hypothetical protein	2
<i>yeaR</i>				-1.11	Hypothetical protein	0
<i>yebG</i>	0.92	1.3	1.38	1.29	Hypothetical protein (ORF96)	3
<i>yebK</i>			-0.88		Hypothetical protein	1
( <i>yefJ</i> )				-0.9	Hypothetical protein	3
<i>yefM</i>				-0.92	Hypothetical protein	3
<i>yfhP</i>	0.97				Hypothetical protein	0
<i>yfhT</i>	1.1				Hypothetical protein	2
<i>yjfF</i>	<b>0.96</b>				Hypothetical protein (f102)	1
<i>yfiD</i>	<b>-1.8</b>			<b>-0.60</b>	Hypothetical protein	1
<i>ygaF</i>				-1.41	Hypothetical protein	2
<i>ygfM</i>		1.62			Hypothetical protein	0
<i>ygfP</i>		1.38			Hypothetical protein	1
<i>ygjJ</i>	<b>2.31</b>	<b>1.43</b>	<b>0.81</b>		Hypothetical protein	2
<i>yhaD</i>			-1.26		Hypothetical protein (ORF3)	3
<i>yhaG</i>			-1.29		Hypothetical protein	2
<i>yhaU</i>			-1.58		Hypothetical protein	1
<i>yhbC</i>	<b>-0.83</b>				Hypothetical protein	2
<i>yhbE</i>	-1.11				Hypothetical protein (f321)	0
<i>yhcH</i>	-1.86		-2.41		Hypothetical protein (f154)	2
<i>yhcN</i>				-1.17	Hypothetical protein	0
<i>yhgN</i>	<b>-1.14</b>			<b>0.75</b>	Hypothetical protein	1
<i>yhhA</i>				-1.2	Hypothetical protein	0
<i>yhiE</i>		-1.45			Hypothetical protein	2
<i>yhjK</i>	0.92				Hypothetical protein	2
<i>yhjX</i>	<-3.00				Hypothetical protein	1
<i>yidK</i>	2.79		>3.34		Hypothetical protein	3
<i>yigB</i>	<b>0.89</b>				Hypothetical protein	1
<i>yihK</i>			<b>1.1</b>		Hypothetical protein	2
<i>yiiD</i>	<b>0.82</b>				Hypothetical protein (o329)	1
<i>yjhA</i>			<-2.81		Hypothetical protein	3
<i>yjhC</i>			-2		Hypothetical protein	0
<i>yjhT</i>	G2(down)		-2.21		Hypothetical protein	1

Table 1. *cont.*

	Relative log ratio (KK335( <i>dam</i> -16:: <i>kam</i> )/KK46) <sup>a</sup>				Description <sup>b</sup>	No. of GATC <sup>c</sup>
	Aerobic		Low aerobic			
	Log	Stationary	Log	Stationary		
<i>yjil</i>	-1.8				Hypothetical protein	4
<i>ykfF</i>			1.72		Hypothetical protein	1
<i>ykgE</i>	<b>-0.96</b>		<b>-0.79</b>		Hypothetical protein	2
<i>yltH</i>				-1.26	Hypothetical protein	5
<i>ymfI</i>		1.33		1.25	Hypothetical protein	1
<i>ymfJ</i>		1.74		1.71	Hypothetical protein	2
<i>ygeW</i>		1.21			Hypothetical protein	2
<i>ygeY</i>		1.42			Hypothetical protein	0
<i>ygeZ</i>		1.39			Hypothetical protein	2
<i>yqaA</i>	<b>-0.81</b>				Hypothetical protein	2
<i>yqeB</i>		1.52			Hypothetical protein	1
<i>yqfB</i>	<b>-1.26</b>			<b>0.78</b>	Hypothetical protein	1
<i>yrbL</i>				-1.54	Hypothetical protein (o210)	1

a. The expression level was described by relative log ratio (logE value) of KK335(*dam*-16::*kam*)/KK46 (*dam*<sup>-</sup>). G2 represents drastic alteration of expression of a gene whose expression was detected only in KK46 (down) or KK335 (up). When some data in four hybridizations were classified as group 2 and others were classified as group 1, the expression level was described by more than (>) or less than (<) the mean of the relative ratio determined by only part of the hybridization data classified in group 1. The details of group 1 and group 2 were described in *Experimental procedures*.

b. The column for description was described according to SWISSPROT, Genobase database and GenBank.

c. The number of GATC sequences within the upstream 500 bp of each gene is described in column c.

The results obtained are representative of four independent hybridizations from two independent total RNAs extracted from independent culture. Only genes with a *P*-value <1E-2 in a Wilcoxon statistical test and factor >0.69 (= twofold alteration corresponding to logE value) differential expression between KK46 (*dam*<sup>-</sup>) and KK335 (*dam*-16::*kam*) (bold), or consistent up/down regulation more than twofold of all four independent hybridizations were classified according to their function.

(*gpsA*) biosynthesis was also increased, as was transcription of genes concerned with purine salvage and interconversion (*apt*, *gpt*, *gst* and *guaC*) and pyrimidine or purine biosynthesis (*adk*, *codA*, *ndk*, *prsA* and *pyrC*). Thus, Dam appears to be involved in the transcriptional regulation of many metabolic pathways.

Notably, in log phase mutant bacteria grown under aerobic and/or low aerobic conditions, transcription of several genes involved in purine nucleotide synthesis was increased. These changes included *purEK*, *purHD*, *purB*, *purF* and *purT*. All these *pur* genes contain common operator sequences (ACGCAAACGTTTGCGT) and are included in the *pur* regulon that is regulated by PurR. However, although the transcription of these individual *pur* genes was increased in the mutant bacteria, the transcriptional levels of *purR* did not differ between the *dam* mutant and the wild-type cells grown in equivalent conditions (data not shown). Thus, how Dam regulates the genes contained in the *pur* regulon is unclear.

#### *The dam* mutation constitutively affects the transcription of genes involved in environmental stress response

The transcriptional activity of genes encoding the heat shock proteins and carrying an *rpoH* (which encodes a heat shock-specific sigma factor in *E. coli*)-dependent promoter (*dnaK*, *dnaJ*, *hslU*, *hslV*, *htpG*, *ibpA*, *groES* and *groEL*, Table 1 and our web site) were all upregulated in

the *dam* mutant grown in aerobic conditions. However, the transcriptome analysis revealed no significant differences in *rpoH* expression levels between the two strains (data not shown). This suggests that some significant stress suffered by the *dam* mutant has led to the accumulation of a heat shock sigma factor, sigma 32, by transcriptional activation or stabilization. The transcript levels of several other environmental response genes (*cspE*, *dppA*, *lytB*, *rnb*, *sfmC* and *surA*) were also increased in the *dam* mutant. It is known that the *dam* mutant has an elevated expression of SOS-related genes such as *recA*, *lexA*, *uvrAB*, *uvrD*, *sulA*, *dinF* and *dinD* (Peterson *et al.*, 1985), and we also observed an increase in the transcription levels of many of these genes (*dinI*, *dinF*, *dinD*, *lexA*, *recA*, *recN*, *ruvA*, *sulA*, *uvrA* and *yebG*; Oh and Kim, 1999) in the *dam* mutant under both aerobic and low aerobic conditions (Table 1 and our web site). Why the *dam* mutant appears to be responding to some constitutively present stress, regardless of environmental conditions, is not clear. One possibility is that GATC sequences are targets for restriction and that, in the Dam mutant, lack of methylation of these sequences might expose GATC-bearing sequences to abnormal restriction (Marinus, 2000). This may in turn lead to a constitutive induction of the SOS response in the *dam* mutant that bears no relation to the environmental conditions. Our data are consistent with this model.

Recently, the transcriptional profile of *E. coli* cells

treated with hydrogen peroxide was examined by DNA microarray analysis (Zheng *et al.*, 2001). Hydrogen peroxide was found to induce the transcription, in an OxyR-independent manner, of a number of genes, including heat shock genes (*groEL*, *groES*, *dnaK* and *htpG*), SOS response genes (*recA*, *recN*, *lexA* and *dinD*), a TCA cycle-related gene (*fumA*), a cysteine metabolism gene (*cysK*) and the *nrd* operon (*nrdF*). In contrast, the expression of many ribosomal protein genes was repressed. With regard to the *dam* mutant, when the cells were grown under aerobic conditions and harvested during the log phase, the transcription of the genes identified by Zheng *et al.* (2001) followed a similar pattern to that for the hydrogen peroxide-treated *E. coli* (Table 1 and our web site). This also suggests that the *dam* mutant suffers some kind of constitutive stress, even when it is grown under favourable aerobic conditions.

*The dam mutation affects the transcriptional expression of genes involved in periplasmic binding protein-dependent transporters*

The transcript level of the *dppBCEF* operon, whose products form the periplasmic-binding protein-dependent (BPD) transport system, is increased in the *dam* mutant under aerobic conditions. Under low aerobic conditions, the expression levels of genes involved in the maltose BPD transport system (the *malEFG* and the *malK-lamB-malM* operons) and involved in change to glucose and glucose 1-phosphate from maltose and maltodextrin (the *malPQ* operon and *malS*) were decreased in the *dam* mutant (Table 1). The transport system is essential to the utilization of maltose and maltodextrins in *E. coli*. We found that the intergenic region between the divergently transcribed *malK* and *malE*, *F* and *G* genes contains multiple binding sites for CRP and MalT overlapping with Dam methylation sites (Fig. 4).

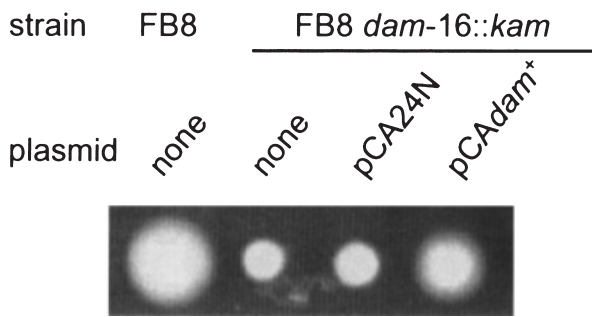
The expression level of genes involved in carbohydrate phosphotransferase systems [*manXYZ* and *nagE*, which are negatively regulated by CRP and NagC (Plumbridge and Kolb, 1991) and *gsr*, *treB* and *ptsH*, which are positively controlled by CRP (De Reuse and Danchin, 1991)] were also decreased in the *dam* mutant under low aerobic conditions (Table 1). In addition, under aerobic and/or low aerobic conditions, genes concerned with the uptake and metabolism of iron, molybdenum and nickel (*fnt*, *modA* and *nika*) were downregulated in the *dam* mutant. Intracellular transport genes (*argT*, *focA*, *nanT*, *nark* and *yaaA*) were also downregulated in the *dam* mutant. The transcription of both *focA* and *nika* is positively regulated by Fnr. Our observations suggest that these genes are regulated by global regulators (CRP, Fnr, MalT and NagC) and Dam via the methylation state of the regulatory region.

*Dam regulates the respiratory- and motility-associated gene expression that allows E. coli to adapt to low aerobic conditions*

As described above, the expression of respiratory genes, especially nitrate reductase, was markedly reduced in the *dam* mutant. Thus, these enzymes may be positively controlled directly or indirectly at the transcriptional level by Dam methylase. Our observations also suggest the intriguing possibility that the methylation of GATC sequences is an integral part of the system used by *E. coli* to adapt to environmental changes, in particular to changes in the levels of oxygen and nitrate, which are the final acceptors in the respiratory chain.

One normal response made by *E. coli* to low oxygen levels is taxis. Taxis is the movement of *E. coli* cells growing in a limited supply of oxygen and various ligands to an environment that has more optimal oxygen and ligand concentrations. The low oxygen concentrations elicit the expression of motility and chemotaxis genes (Jones *et al.*, 1992) when methyl-accepting chemotaxis proteins (Aer, Tar, Tsr, Tap and Trg) sense signals that consist of oxygen levels, redox potential, light and external ligands. These signals are then transmitted to cytoplasmic signal transduction proteins (CheA, CheB, CheR, CheW, CheY and CheZ) and affect the motility of *E. coli* by inducing taxis (Taylor and Zhulin, 1998; Taylor *et al.*, 1999). Alternative acceptors in the respiratory chain, such as nitrate, can also elicit an aerotaxis-like behaviour by substituting for oxygen as an electron acceptor (Taylor *et al.*, 1979; Taylor and Zhulin, 1998). Flagellar, motility and chemotaxis genes are known to form clusters in four distinct regions on the chromosome, namely in regions I (24 min), II (41 min), IIIa and IIIb (43 min) (Macnab, 1996). When the *dam* strain was grown under low aerobic conditions, its expression of most genes located specifically in region II (*tar-tap-cheRBYZ* operon, *motAB-cheAW* operon) and region IIIa (*fliC*) was significantly lower than in the wild-type cells (Table 1). It suggests that Dam is involved in regulating taxis under low aerobic conditions and that the *dam* mutant might be defective in taxis under these conditions.

These changes in the *dam* mutant cell suggest that its motility may differ from that of the wild-type strain, and thus we tested the motility of both strains. The wild-type strain (FB8) shows normal motility, but the isogenic *dam* mutant (FB8 *dam-16::kam*) strain did not move normally (Fig. 1). Confirming that *dam* is involved, the *dam* mutant complemented by plasmid pCA*dam*<sup>+</sup> showed normal motility (Fig. 1), whereas the *dam* mutant carrying the pCA24N plasmid, which is the vector system used for archiving all the *E. coli* ORFs (Mori *et al.*, 2000), did not move normally. These results were also observed in the MG1655, KK46 (W3110 derivative strain) and their iso-



**Fig. 1.** Motility of the wild-type strain FB8, the FB8 *dam-16::kam* mutant, FB8 *dam-16::kam* containing plasmid pCA24N and FB8 *dam-16::kam* containing plasmid pCAdam<sup>+</sup>. Fresh overnight cultures of each strain (1  $\mu$ l) were spotted onto a semi-solid agar plate, incubated at 30°C for 6 h and photographed. The segregation frequency, whereby the plasmid-bearing cells lost the plasmid during growth, was very low, as the frequency of plasmid-loss cells among the total population was only <0.5% after 50 generations.

genic *dam* mutant strains (data not shown), although the native motility of KK46 was less than that of the FB8 and MG1655 strains. Thus, loss of *dam* activity generated a defective motility phenotype.

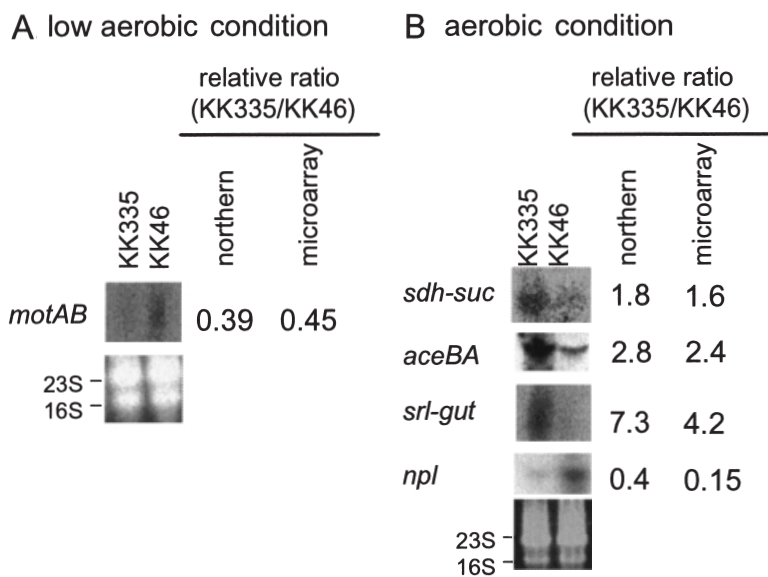
#### Poor motility may be responsible for the lack of invasiveness by the *S. typhimurium dam* mutant

The *dam* mutant of *S. typhimurium* was recently reported to be defective in cell invasion and virulence (Garcia Del Portillo *et al.*, 1999; Heithoff *et al.*, 1999), indicating the importance of Dam in regulating these activities. However, the crucial virulence/invasion genes being modulated by Dam in *S. typhimurium* were not identified. As motility-negative mutant strains of *S. typhimurium* are also non-invasive (Liu *et al.*, 1988; Jones *et al.*, 1992), it may

be that Dam-regulated motility genes are important in virulence/invasion. This notion is supported by our parallel observations made with *E. coli*. We showed that, in the *dam* mutant of *E. coli* grown under low aerobic conditions, the expression of certain chemotaxis-related genes (*tar*, *tap*, *cheA*, *cheY*, *motA* and *motB*) is significantly lower than in the wild type (Table 1). The motility of the *dam* mutant was also less than that of wild type (Fig. 1). Given the close sequence similarities between *S. typhimurium* and *E. coli*, it is quite possible that the *dam*-specific expression profile of the *E. coli dam* mutant may also occur in the *dam* mutant of *S. typhimurium*, supporting the notion that poor motility may be responsible for the poor invasiveness of the *dam* mutant of *S. typhimurium*.

#### Results obtained with Northern hybridization and DNA microarray methods are comparable

We assessed whether the results obtained with the DNA microarray method could be reproduced using another method for detecting mRNA, i.e. Northern hybridization. Thus, the wild-type and the *dam* mutant strains were cultured under aerobic and low aerobic conditions and harvested at log phase. Total RNA was prepared from the cell extract of each strain and separated by gel electrophoresis. The blotted membrane was then hybridized with several labelled probes. For examining gene expression in aerobic conditions, DNAs of the *sdhC* (the *sdh-suc* operon), *aceB* (the *aceBA* operon), *gutM* (the *srl-gut* operon) and *npl* were amplified by polymerase chain reaction (PCR) and labelled with [<sup>32</sup>P]-dCTP. For examining gene expression in low aerobic conditions, *motA* (the *motAB* operon) probes were made and labelled. The hybridized membranes are shown in Fig. 2 and indicate



**Fig. 2.** Comparison of the mRNA levels of five genes in the KK46 and KK335 strains measured by Northern blot and microarray analyses.

A. Northern blot analysis of the *motAB* operon was performed using RNA prepared twice from KK46 and KK335 (*dam-16::kam*) grown in low aerobic conditions. The KK335/KK46 mRNA ratios for both genes were calculated from data obtained by the Northern blot and microarray analyses and are shown to the left of the data. The same RNA preparations were used for both methods.

B. Northern blot analysis of the *sdh-suc* operon, the *aceBA* operon, the *srl-gut* operon and *npl* was performed using RNA prepared from KK46 and KK335 (*dam-16::kam*) grown in aerobic conditions. The KK335/KK46 mRNA ratios were calculated as in (A) and are shown to the left of the data. The relative ratios of microarray data for each gene were taken from [http://ecoli.aist-nara.ac.jp/xp\\_analysis/dam/all.html](http://ecoli.aist-nara.ac.jp/xp_analysis/dam/all.html). Each relative ratio represents the mean of four independent microarray data.

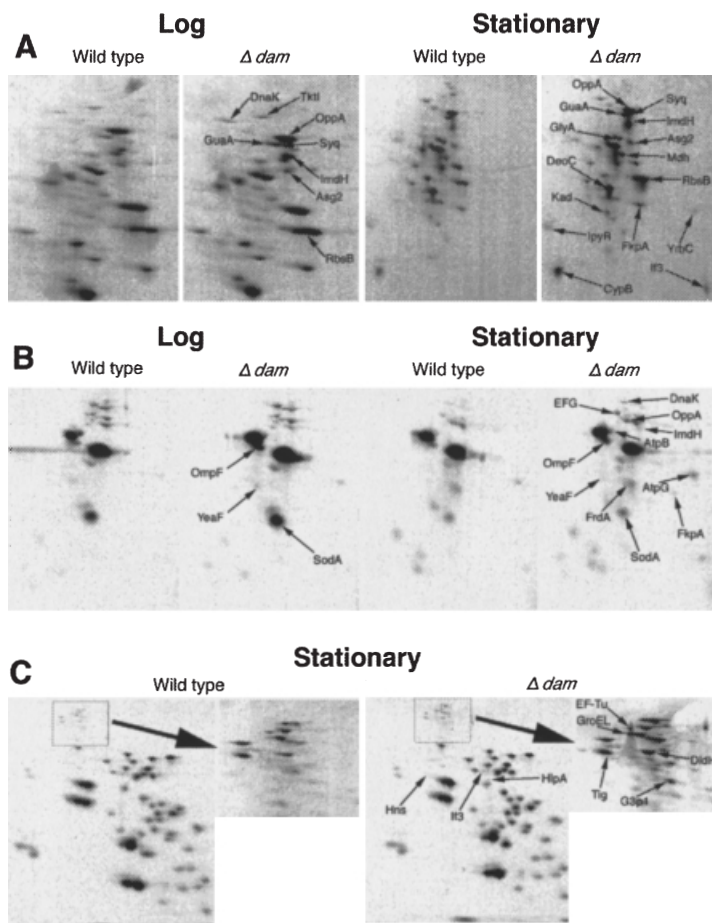
that the transcription levels of these genes as measured by the DNA microarray method are consistent with those measured by Northern hybridization.

#### Two-dimensional gel electrophoresis of proteins from the $\Delta dam$ mutant supports the microarray analysis

To examine the expression levels of the proteins encoded by the genes whose transcription was significantly altered in the  $dam$  mutant, we carried out proteome analysis by radical-free and highly reducing two-dimensional polyacrylamide gel electrophoresis (RFHR 2-D PAGE) (Wada, 1986). We examined the proteins fractionated from wild-type and  $dam$  mutant cells growing at log and stationary phases under aerobic and low aerobic conditions. As expected from the microarray assays, products of genes involved in a number of cellular functions were present at higher levels in the  $dam$  strain grown under aerobic conditions compared with similarly grown wild-type cells (Table 2A and B, Fig. 3). These categories included energy metabolism (*atpG*, *frdA*, *gapA*, *lpdA*, *mdh*), translation (*fusA*, *glnS*, *infC*, *ppiB*, *tufA*), cell envelope (*ompF*, *yeaF*, *uncD*), amino acid metabolism (*ansB*, *glyA*),

nucleotide metabolism (*adk*, *deoC*, *guaA*, *guaB*) and stress response (*dnaK*, *fkpA*, *groEL*, *tig*). In addition, the accumulation of the oxidative stress-responsive gene product SodA, the nucleoid protein H-NS and HlpA and the OppA, RbsB, lpyR and YrbC proteins were clearly observed in the  $dam$  mutant at the stationary phase. These data strongly suggest that the expression of TCA cycle-related enzymes, as well as GuaA, lmdH and DeoC, which participate in the purine and pyrimidine salvage pathway, is modulated by Dam (see microarray results described above). This analysis also revealed that gene products involved in the translational machinery and the chaperone system for protein folding and transport are more highly expressed in the  $dam$  mutant than in the wild type grown under aerobic conditions (Table 2A and B, Fig. 3). On the other hand, EF-G, GatY and ModA in the log phase and OmpW in the stationary phase were decreased in the  $dam$  mutant, although we could find no biological meaning for this downregulation at present (Table 2A and B).

When the  $dam$  mutant was grown under low aerobic conditions and harvested at log phase, the levels of proteins encoded by genes involved in energy metabolism



**Fig. 3.** RFHR 2-D PAGE (proteome) analysis of proteins in KK46 and KK335 (*dam-16::kam*) grown to the log or stationary phase. The bacterial samples from KK46 and KK335 (*dam-16::kam*) grown in L broth under aerobic conditions were fractionated and analysed by two-dimensional electrophoresis. The two-dimensional patterns with PRS (post-ribosomal supernatant), CD (crude debris) and CR (crude ribosome) fractions are presented in (A), (B) and (C) respectively. Arrows indicate proteins that were detected in increased quantities (ratio >1.3) in KK335 (*dam-16::kam*) compared with the wild type. A summary of this analysis is presented in Table 2A and B. There was no significant difference between KK335 (*dam-16::kam*) and the wild-type strains in the two-dimensional pattern of CR at log phase (data not shown). The two-dimensional experiments were repeated at least three times, and these patterns were confirmed to be reproducible. The data from one such representative experiment are shown.

Table 2. Summary of proteome analyses of the *dam* mutant strain under aerobic and low aerobic conditions.

Protein	Gene	Description	Functional classification <sup>a</sup>	Protein <sup>b</sup> KK335/KK46	cDNA <sup>c</sup> KK335/KK46	No. of <sup>d</sup> GATC sites
<b>A. Log phase (aerobic condition)</b>						
Asg2	<i>ansB</i>	Asparaginase II precursor	Amino acid metabolism	1.4	0.4	2
OmpF	<i>ompF</i>	Outer membrane protein F precursor	Cell envelope	1.5	0.8	1
YeaF	<i>yeaF</i>	Scarfolding protein for murein-synthesizing holoenzyme	Cell envelope	1.3	1.9	1
GatY	<i>gatY</i>	Tagatose-bisphosphate aldolase	Energy metabolism	0.7	0.4	1
SodA	<i>sodA</i>	Superoxide dismutase	Environmental response	1.3	1	1
GuaA	<i>guaA</i>	GMP synthase	Nucleotide metabolism	1.4	1.3	1
ImdH	<i>imdH</i>	IMP dehydrogenase	Nucleotide metabolism	1.5	1.1	3
OppA	<i>oppA</i>	Periplasmic oligopeptide-binding protein	Protein/peptide secretion	1.4	1.3	1
EF-G	<i>fusA</i>	Elongation factor G	Translation	0.4	0.8	1
Syq	<i>glnS</i>	Glutaminyl-tRNA synthetase	Translation	1.4	1.8	5
ModA	<i>mocA</i>	Molybdate-binding periplasmic protein precursor	Transport/binding protein	0.7	0.5	0
RbsB	<i>rbsB</i>	D-ribose-binding periplasmic protein precursor	Transport/binding protein	1.3	1.6	4
<b>B. Stationary phase (aerobic condition)</b>						
Asg2	<i>ansB</i>	Asparaginase II precursor	Amino acid metabolism	1.4	1.3	2
GlyA	<i>glyA</i>	Serine hydroxymethyltransferase	Amino acid metabolism	1.4	1	1
AtpB	<i>uncD</i>	H <sup>+</sup> transporting ATP synthase beta chain	Cell envelope	1.8	1.3	4
OmpW	<i>ompW</i>	Outer membrane protein W	Cell envelope	0.7	0.6	3
YeaF	<i>yeaF</i>	Scarfolding protein for murein-synthesizing holoenzyme	Cell envelope	5.8	0.9	1
IpyR	<i>ppa</i>	Inorganic pyrophosphatase	Central intermediary metabolism	1.4	0.8	1
AptG	<i>aptG</i>	ATP synthase gamma chain	Energy metabolism	2.5	1.6	3
DidH	<i>lpdA</i>	Dihydroipoamide dehydrogenase	Energy metabolism	2.8	1.9	1
FrdA	<i>frdA</i>	Fumarate reductase flavoprotein subunit	Energy metabolism	1.3	1.4	4
G3P1	<i>gapA</i>	Glyceraldehyde-3-phosphate dehydrogenase A	Energy metabolism	3.5	0.9	2
Mdh	<i>mdh</i>	Malate dehydrogenase	Energy metabolism	1.3	1.2	0
DnaK	<i>dnaK</i>	DnaK protein	Environmental response	1.9	0.9	1
FkpA	<i>fkpA</i>	Fkbp-type peptidyl-prolyl <i>cis-trans</i> isomerase	Environmental response	1.7	1.3	2
GroEL	<i>groEL</i>	GroEL protein	Environmental response	3.9	1.2	5
SodA	<i>sodA</i>	Superoxide dismutase	Environmental response	1.3	0.8	1
Tig	<i>tig</i>	Trigger factor	Environmental response	1.7	1.2	0
DeoC	<i>deoC</i>	Deoxyribose-phosphate aldolase	Nucleotide metabolism	1.3	1.6	2
GuaA	<i>guaA</i>	GMP synthase	Nucleotide metabolism	1.7	0.9	1

Table 2. cont.

Protein	Gene	Description	Functional classification <sup>a</sup>	Protein <sup>b</sup> KK335/KK46	cDNA <sup>c</sup> KK335/KK46	No. of <sup>d</sup> GATC sites
ImdH	<i>guaB</i>	Imp dehydrogenase	Nucleotide metabolism	1.9	0.8	3
Kad	<i>adk</i>	Adenylyate kinase	Nucleotide metabolism	1.4	1.9	4
OppA	<i>oppA</i>	Periplasmic oligopeptide-binding protein	Protein/peptide secretion	1.5	1.2	1
HlpA	<i>hlpA</i>	Histone-like protein Hlp-1 precursor	Regulatory function	1.5	1.5	3
H-NS	<i>hns</i>	DNA binding protein	Regulatory function	1.8	0.7	0
CypB	<i>pppB</i>	Peptidylprolyl isomerase	Translation	1.7	1.2	5
EF-G	<i>fusA</i>	Elongation factor G	Translation	2.4	0.9	1
EF-TU	<i>tufA</i>	Translation elongation factor EF-TU.A	Translation	2.4	1.5	2
If3	<i>infC</i>	Initiation factor IF-3	Translation	1.4	0.6	1
Syq	<i>glnS</i>	Glutaminyl-tRNA synthetase	Translation	1.7	1.3	5
RbsB	<i>rbsB</i>	D-ribose-binding periplasmic protein precursor	Protein/peptide secretion	1.9	1.5	4
YrbC	<i>yrbC</i>	Hypothetical protein	Unknown	1.5	0.9	1
C. Log phase (low aerobic condition)						
Otc2	<i>argF</i>	Ornithine carboxyltransferase (EC 2.1.3.3) chain F	Amino acid metabolism	2.4	1.2	1
GlyA	<i>glyA</i>	Serine hydroxymethyltransferase	Amino acid metabolism	1.7	1.2	1
IpyR	<i>ppa</i>	Inorganic pyrophosphatase	Central intermediary metabolism	3.8	1	1
AtpB	<i>atpD</i>	H <sup>+</sup> transporting ATP synthase beta chain	Cell envelope	1.4	1.3	4
OmpW	<i>ompW</i>	Outer membrane protein W	Cell envelope	1.3	0.7	3
Alf	<i>fab</i>	Fructose-bisphosphate aldolase	Energy metabolism	3.1	0.7	0
FrdA	<i>frdA</i>	Fumalate reductase	Energy metabolism	0.4	1	4
G3pl	<i>gapA</i>	Glyceraldehyde-3-phosphate dehydrogenase	Energy metabolism	1.6	0.7	2
GatY	<i>gatY</i>	Tagatose-bisphosphate aldolase	Energy metabolism	0.6	0.2	1
GlpD	<i>glpD</i>	Glycerol-3-phosphate dehydrogenase	Energy metabolism	2.9	1.1	2
Pgk	<i>pgk</i>	Phosphoglycerate kinase	Energy metabolism	3.0	0.7	4
TpiS	<i>tpiA</i>	Triosephosphate isomerase	Energy metabolism	2.2	0.9	1
DegP	<i>htrA</i>	Heat shock protein protease Do precursor	Environmental response	0.4	0.6	0
DnaK	<i>dnaK</i>	DnaK protein, chaperone Hsp70	Environmental response	2.9	0.7	1
DppA	<i>dppA</i>	Dipeptide-binding protein	Environmental response	2.9	0.7	1
GroEL	<i>groEL</i>	Chaperonin Hsp60	Environmental response	3.8	0.6	5
HlpA	<i>hlpA</i>	Histone-like protein	Regulatory function	1.6	0.9	3
DeaD	<i>deaD</i>	ATP-dependent RNA helicase	Transcription	3.9	2.7	1
RhIE	<i>rhIE</i>	Putative ATP-dependent RNA helicase RhIE	Transcription	1.9	5.8	0
Rho	<i>rho</i>	Transcription termination factor G	Transcription	1.7	0.6	2



EF-G	<i>fusA</i>	Elongation factor G	Translation	4.1	1.4	1
RL10	<i>rplJ</i>	Ribosomal protein L10	Translation	1.8	1.5	0
Rrf	<i>rff</i>	Ribosomal recycling factor	Translation	1.6	1.2	2
RS1	<i>rpsA</i>	30S ribosomal protein S1	Translation	2.9	1.9	2
RS3	<i>rpsC</i>	30S ribosomal protein S3	Translation	2.3	1.6	3
OppA	<i>oppA</i>	Periplasmic oligopeptide-binding protein	Transport/binding protein	1.6	1.0	1
ProP	<i>proP</i>	Proline/betaine transport protein	Transport/binding protein	1.9	1.4	1
RbsB	<i>rbsB</i>	D-ribose-binding periplasmic protein precursor	Transport/binding protein	1.8	1.4	1
YajQ	<i>yajQ</i>	Hypothetical protein	Unknown	1.3	1.4	4
YbeJ	<i>ybeJ</i>	Amino-acid ABC transporter binding protein	Unknown	2.5	0.9	2
YdgH	<i>ydgH</i>	Hypothetical protein	Unknown	1.8	1.1	1
D. Stationary phase (low aerobic condition)						
MotA	<i>motA</i>	Chemotaxis protein MotA	Cell envelope	0.4	1.7	1
OmpC	<i>ompC</i>	Outer membrane c precursor	Cell envelope	3.2	0.6	0
Alf	<i>fbA</i>	Fructose-bisphosphate aldolase	Energy metabolism	0.3	0.7	0
Eno	<i>eno</i>	Enolase	Energy metabolism	0.5	0.6	3
G3P1	<i>gapA</i>	Glyceraldehyde-3-phosphate dehydrogenase A	Energy metabolism	0.3	0.6	2
GatY	<i>gatY</i>	Tagatose-bisphosphate aldolase	Energy metabolism	0.2	0.5	1
ODP1	<i>ace</i>	Pyruvate dehydrogenase el component	Energy metabolism	0.4	0.9	0
PflB	<i>pflB</i>	Formate c-acetyltransferase	Energy metabolism	0.4	0.5	2
Pgk	<i>pgk</i>	Phosphoglycerate kinase	Energy metabolism	0.1	0.9	4
TpiS	<i>tpiA</i>	Triosephosphate isomerase	Energy metabolism	0.7	1.0	1
DnaK	<i>dnaK</i>	DnaK protein, chaperone Hsp70	Environmental response	0.5	0.4	1
DnaJ	<i>dnaJ</i>	DnaJ protein, chaperone Hsp40	Environmental response	0.5	0.7	0
GroEL	<i>groEL</i>	Chaperonin Hsp60	Environmental response	0.7	0.8	5
SodA	<i>sodA</i>	Superoxide dismutase	Regulatory function	0.2	0.6	1
Dps	<i>dps</i>	DNA binding protein	Regulatory function	0.5	0.5	0
RS1	<i>rpsA</i>	30S ribosomal protein S1	Translation	1.5	0.5	2
OppA	<i>oppA</i>	Periplasmic oligopeptide-binding protein	Transport/binding protein	1.6	0.8	1
PthP	<i>pthP</i>	Phosphocarrier protein Hpr	Transport/binding protein	0.6	0.5	1
SR54	<i>ffh</i>	Signal recognition particle protein	Transport/binding protein	0.2	0.7	1
YajQ	<i>yajQ</i>	Hypothetical	Unknown	1.7	0.6	4

a. The proteins are classified into biological functional categories modified from Riley and Labedan, 1996).

b. Ratios of amounts of protein, KK335(*dam-16::kam*)/KK46(*dam*<sup>+</sup>), were analysed by RFHR 2-D PAGE.

c. Ratios of amounts of cDNA, KK335(*dam-16::kam*)/KK46(*dam*<sup>+</sup>), were analysed by DNA microarrays (see *Experimental procedures* for microarray).

d. Number of GATC sequences on 500bp upstream region from the first base of ORF start site. Proteins whose expression was significantly enhanced or decreased in the *dam* mutant are listed: ratios of amounts of protein, KK335 (*dam-16::kam*)/KK46(*dam*<sup>+</sup>), >1.3 or <0.7 at log and stationary phase are listed. The ratios obtained by DNA microarray analysis under the same conditions are shown for comparison.

(*fba*, *gapA*, *glpD*, *tpiA*), stress response (*dnaK*, *dppA*, *groEL*), transcription (*deaD*, *rhIE*, *rho*), translation (*fusA*, *rplJ*, *rrf*, *rpsA*, *rpsC*) and transport/binding protein (*proP*, *rbsB*, *oppA*, *ybeJ*) were increased (Table 2C). On the other hand, when *dam* cells were grown in low aerobic conditions and harvested in the stationary phase, the protein products of genes involved in energy metabolism (*tpiA*, *gapA*, *pgk*, *eno*, *fba*, *gatY*, *ace*, *pfIB*) were decreased. Similarly, proteins involved in stress response (*dnaK*, *dnaJ*, *groEL*, *sodA*), the transport/binding protein (*ptsH*, *ffh*) and the DNA-binding protein (*dps*) were also present in lower amounts in the *dam* mutant (Table 2D). The levels of many ribosomal proteins were also decreased (data not shown).

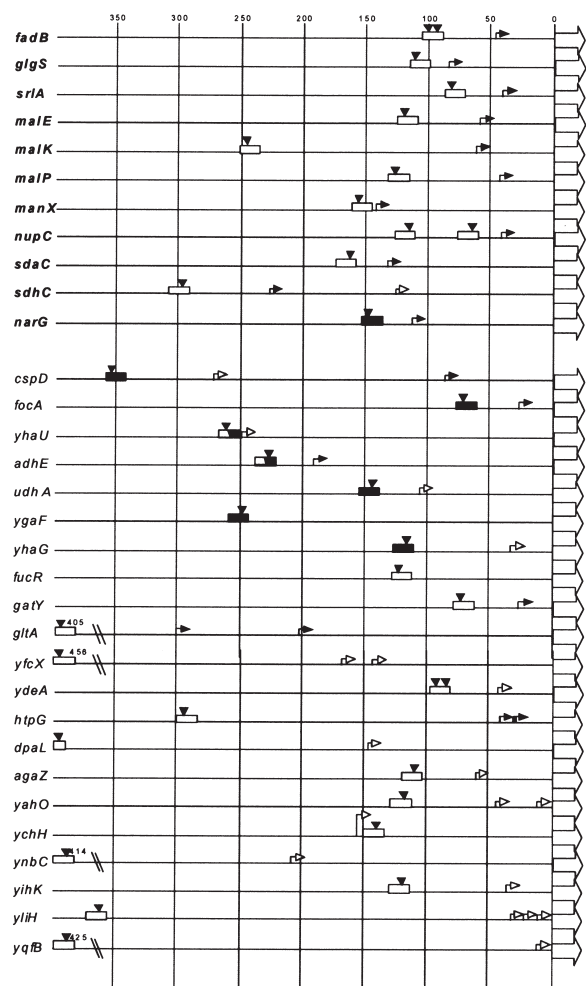
In general, the observations made with the two-dimensional analysis were consistent with those arising from the microarray analysis.

#### GATC sites and the cis-elements of global regulators overlap in the promoters of genes affected by Dam deficiency

The previously published computational analysis of the GATC sequences in part of the *E. coli* genome revealed that many GATC sequences overlap with consensus Fnr and/or CRP binding sites (Henaut *et al.*, 1996). We thus examined the genes affected by the *dam* mutation that we identified in our microarray analysis for the presence of GATC sites that overlap with documented or predicted Fnr and/or CRP binding sites. In many cases, an overlap was observed, indicating that the transcription of these genes may be directly regulated by Dam-mediated DNA methylation that modulates the activity of transcriptional regulators (Fig. 4). However, GATCs were not found within the recognition sequences for transcriptional regulators in some cases, suggesting that *dam*-mediated regulation of some genes in *E. coli* may also occur through indirect mechanisms.

#### Conclusions

Dam methylase is found in a limited number of Proteobacteria species, namely the gamma subdivision, which includes *E. coli* and *S. typhimurium*. Dam-mediated methylation is known to be responsible for regulating metabolism, invasion, replication and mismatch repair in *E. coli* (Marinus, 1996; 2000; Garcia Del Portillo *et al.*, 1999; Heithoff *et al.*, 1999). In addition, DNA methylation of the *oriC* region by Dam is indispensable for the regulation of DNA replication and the organization of daughter chromosome separation. Interestingly, two other proteins that, like Dam, exist only in *E. coli* and related bacteria, also play an important role in the positioning of daughter chromosomes, namely SeqA, which binds to



**Fig. 4.** Genes found to be differentially regulated in KK46 and KK335 and that have a documented or putative CRP or Fnr binding site that overlaps with GATC sequences. ORFs of genes (operons) are represented as boxed arrows. The 400 bp upstream of each ORF are represented by a straight line. Intervals of 50 bp are indicated by vertical lines. Black arrows identify the documented transcriptional start sites. Open arrows identify the putative transcriptional start sites (data obtained from NCBI). Open boxes indicate documented or putative CRP binding sites (GTGANNNNNTCAC, allowing for a 1 bp mismatch). Filled boxes indicate documented or putative Fnr binding sites (TTGANNNNNTCAA; allowing for a 1 bp mismatch). Open/filled boxes indicate putative CRP and/or Fnr binding sites. Arrowheads on the open or closed boxes indicate GATC sites within the CRP or Fnr binding sites respectively. CRP or Fnr binding sites of genes written in bold letters have been documented to be CRP or Fnr binding sites (Li *et al.*, 1994; Basby and Kolb, 1996). If recognition sequences were predicted to be present in the 400–500 bp upstream, the 3' position of these recognition sequences is shown at the relevant bp number.

hemimethylated GATC sequences, and MukB (Hiraga *et al.*, 2000). The methylation of GATC by Dam may thus constitute a unique and important system that regulates biological function in some bacteria.

Our DNA microarray and two-dimensional electrophoretic analyses suggest that energy metabolism and

the stress response are activated inappropriately in the *dam*-deficient mutant grown in aerobic conditions. These findings strongly suggest that at least one of the functions of Dam is the negative regulation of energy metabolism under aerobic conditions. In contrast, the expression of taxis-related and nitrate reductase genes was decreased in the *dam* mutant grown in low aerobic conditions. It thus appears that energy synthesis in the *dam* mutant is insufficient, especially in low aerobic conditions. These observations together suggest that Dam might properly regulate the balance of energy synthesis according to the aerobic/anaerobic conditions.

We also observed that transcription of motility genes was poor in the *dam* mutant, and functional studies confirmed that this mutant indeed had poor motility. Given that motility-deficient and *dam*-deficient *S. typhimurium* mutants are both poorly invasive, our observations with the closely related *E. coli* support the notion that the low invasiveness of the *S. typhimurium dam* mutant may result from its poor motility.

One mechanism by which Dam regulates the transcription of the genes affected in the *dam* mutant is by methylating sites within transcriptional regulators, thus modulating their activity (Fig. 4). However, as GATC sequences were not found to coincide with transcriptional regulator recognition sequences in the promoter regions of some genes, it appears that *dam*-mediated regulation may also occur through indirect mechanisms.

That the regulation of fundamental cellular activities such as energy metabolism and the oxidative stress response is Dam dependent suggests that further analysis of the mechanism by which Dam regulates transcription will provide important clues about how *E. coli* and related bacteria respond to environmental change.

## Experimental procedures

### Bacterial strains, plasmids, media and growth conditions

Isogenic bacterial strains derived from the *E. coli* K-12 strain were used. The wild-type strain is denoted as KK46 (YK1100: *trpC9941*) (Yamanaka *et al.*, 1996), and the *dam*-defective mutant is KK335 (KK46: *dam-16::kam*). The amount of Dam methyltransferase (MTase) in the mutant *dam-16::kam* is below the level of detection both *in vitro* and *in vivo* (Parker and Marinus, 1988). KK335 is a deletion–insertion mutant that was derived by P1 transduction from KK46 (S. Hiraga, unpublished data). Two other *E. coli* strains denoted as FB8 and MG1655 were used for motility assays (see below). Their *dam-16::kam*-deficient isogenic mutant strains were also constructed by P1 transduction. Several plasmids, namely pCA24N and pCAdam<sup>+</sup>, were used for complementation analysis of the low motility phenotype of the *dam* mutant FB8 and MG1655 strains. pCAdam<sup>+</sup> was constructed by *NotI* digestion and self-ligation from the *dam*<sup>+</sup> archive clones (Mori *et al.*, 2000), and its sequence was confirmed by DNA sequencing analysis. All cultures were grown in Luria–Bertani

(LB) medium. Overnight cultures were diluted 500-fold with fresh LB medium and cultured further. Cells were grown to the log (OD<sub>600</sub> = 0.4) or stationary phase (time 5–9 h) in either aerobic (5 h) or low aerobic conditions (9 h). Four independent cultures for each growth phase–aerobicity combination were harvested. Aerobic cultures were grown at 37°C in 1 l flasks containing 200 ml of medium that was reciprocally shaken at 170 r.p.m. Low aerobic cultures were grown in the same conditions except that they were in a 300 ml flask containing 200 ml of LB broth and under rotary shaking. The concentration of dissolved oxygen (DO) in the culture medium was monitored during cell culture using an OM-12 DO meter (Horiba). The DO (3.2 mg l<sup>-1</sup>) of cultures grown to log phase under aerobic conditions was about five times higher than that in the log phase under low aerobic conditions (DO 0.63 mg l<sup>-1</sup>). However, the DO (0.23 mg l<sup>-1</sup>) of the aerobic culture grown to stationary phase was very similar to the DO (0.29 mg l<sup>-1</sup>) of cultures grown to stationary phase in low aerobic conditions.

### RNA isolation

Preparation of total RNA from crude cell lysate was performed using a modified hot phenol method (Aiba, 1985). Briefly, cells were harvested by centrifugation at 12 000 g for 2 min at room temperature, resuspended in 0.5 ml of solution A (0.5% SDS, 20 mM sodium acetate, 10 mM EDTA) and then mixed by pipetting with 0.5 ml of acidic phenol (pH 5.5) preheated at 60°C. The mixture was incubated at 60°C for 5 min. After centrifugation at 12 000 r.p.m. for 3 min at room temperature, the supernatant was recovered. This extraction process was repeated. A phenol–chloroform (1:1, pH 5.5) extraction was then performed, and the RNA was precipitated by the addition of three volumes of ethanol. The RNA pellet was dried and dissolved in a DNase solution (100 mM sodium acetate, 50 mM MgSO<sub>4</sub>) containing 5 units of RNase-free DNase (Takara), and incubated at room temperature for 1 h. A second phenol–chloroform extraction and RNA precipitation were then performed. Purified total RNA was subjected to 1% agarose gel electrophoresis to check for degradation and whether the 23S and 16S ribosomal RNA were recovered without contamination of genomic DNA. The prepared RNAs were used for both microarray and Northern blot analyses.

### Preparation of DNA microarrays

We used custom-made high-density microarrays of DNA molecules on glass slides that had been prepared by the Takara Shuzo Company. The array contains the 4097 independent genes of the *E. coli* genome that have been cloned previously from the *E. coli* K-12 W3110 strain, the so-called archive clone (Mori *et al.*, 2000). Each gene on the slide was completely amplified by PCR using vector-specific primers targeting both sites of the integrated gene fragment: primer 1, 5'-ATCACCAT CACCATACGGATCCGGCCCTGA-3'; primer 2, 5'-TTCTTCT CCTTACTGCGGCCGCATAGGCC-3'.

The PCR-amplified fragments contained the DNA region spanning from the second to the last codon of all genes in *E. coli*. The DNA concentration was more than 0.1 mg ml<sup>-1</sup>. Furthermore, all PCR fragments were confirmed by DNA

sequencing. In addition to the genes mentioned above, there were 24 spots of human  $\beta$ -actin gene as a negative control on the slide. Human transferrin receptor gene, *E. coli* genomic DNA and a fluorescent position marker were spotted as a negative control, a positive control and a positional marker, respectively, to estimate the spotting error. We also checked the accuracy of our microarray using well-characterized mutants. Using four two-component deletion mutants, namely,  $\Delta arcA$ ,  $\Delta arcB$ ,  $\Delta ompR$ –*envZ* and  $\Delta rrsB$ , the DNA microarray analyses indicated, as expected, upregulation of the TCA cycle genes in the  $\Delta arcA$  and  $\Delta arcB$  mutants, downregulation of the *ompFC* transcript in the  $\Delta ompR$ –*envZ* mutant and increased transcription of *tpoS*-dependent genes in the  $\Delta rrsB$  mutant (T. Oshima, H. Aiba, Y. Masuda, S. Kanaya, B. L. Wanner, H. Mori and T. Mizuno, submitted). The accuracy of our microarray was also verified by the fact that our observations regarding SOS gene induction in the *dam* mutant coincided with those reported previously (Peterson *et al.*, 1985). The microarray of the *E. coli* genome is now available from Takara Shuzo.

#### Fluorescent-labelled cDNA preparation, array hybridization and the capture of data

The preparation of fluorescent-labelled cDNA using Cy3 and Cy5, and microarray hybridization were performed essentially according to the M guide (<http://cmgm.stanford.edu/pbrown/mguide/index.html>; DeRisi *et al.*, 1997). Fluorescent-labelled cDNA probes were prepared by random priming methods. Reverse transcriptase reactions were performed by AMVXL (XL Life Science) and 4 nmol of either Cy3-dUTP or Cy5-dUTP (Amersham Pharmacia) using the total RNA from KK46 or KK335 (*dam-16::kam*) grown in the four conditions as templates respectively. Labelled cDNA probes were purified by Centri-sep (Princeton Separations), phenol–chloroform extraction and ethanol precipitation. After drying, the cDNA probe was dissolved in 9  $\mu$ l of water. Both Cy3- and Cy5-labelled cDNA probes were then added to a final volume of 23  $\mu$ l of hybridization buffer (4 $\times$  SSC, 0.2% SDS, 5 $\times$  Denhardt's solution, 100 ng ml<sup>-1</sup> salmon sperm DNA) and denatured by heating at 98°C for 2 min. The denatured cDNA probe was applied to the microarray prehybridized by 100 ng ml<sup>-1</sup> salmon sperm DNA under a coverslip. Hybridization was carried out at 65°C for 16 h. Slides were washed at 60°C with 2 $\times$  SSC for 5 min, then at 60°C with 0.2 $\times$  SSC containing 0.1% SDS and, finally, at room temperature with 0.2 $\times$  SSC. The slides were scanned for fluorescent intensity using a GMS 418 array scanner (Genetic Microsystems) and recorded to 16 bit image files. The signal density of each spot in the microarray was quantified using IMAGE software (BioDiscovery). Two independently obtained mRNA preparations of cells at each growth phase–aerobicity combination were tested. Each preparation was then tested twice by microarray analysis. Thus, the values shown in Table 1 are the mean of the four independently obtained data per gene (spot).

#### Data analysis of microarrays

To distinguish reliable data from the background, we corrected each spot for the local background by subtracting the

local background from the intensity of each spot. In addition, a mean value of the intensity of the 24 negative control spots (human  $\beta$ -actin gene) was determined, together with a standard deviation (SD). Each spot or gene, now represented by a corrected signal intensity, was then classified into three groups according to the relative expression of the gene in the wild type and the *dam* mutant. Group 1 consisted of genes in which both the Cy3 (wild type) and Cy5 (mutant) signal intensities were greater than the mean +1 SD of the negative controls. Group 2 consisted of spots in which either, but not both, the Cy3 and Cy5 signal intensities were greater than the mean +1 SD of the negative controls. Group 3 consisted of spots in which both Cy3 and Cy5 signal intensities were lower than the mean +1 SD of the negative controls. We then normalized the intensity of all spots in group 1. This was done by calculating, for each spot in group 1, the following ratio: mRNA level from KK335 (*dam-16::kam*) labelled by Cy5/mRNA level from KK46 labelled by Cy3. Initially, all group 1 spots were normalized by defining the mean of ratios (Cy5/Cy3) of all spots as 1.0. The ratio of the group 2 spots could not be determined because of the lack of either a Cy3 or a Cy5 fluorescent signal. Spots in this group with high Cy3 or Cy5 intensity (over 1000, i.e. of a sufficiently high intensity value to be detected precisely by the GMS 418 array scanner) were considered to represent altered expression levels in the *dam* mutant relative to the wild-type strain. Spots in group 3 were considered to be undetectable spots. The expression profiles of KK335 (*dam-16::kam*) were compared with those of the KK46 strain in two ways. (i) To ensure that the observed transcript alterations were really caused by the *dam* mutant, we first assessed the degree of random fluctuation and systematic biases inherent in our system. Cy5- and Cy3-labelled probes were simultaneously synthesized from the same template RNA purified from the wild-type strain, and their intensities were compared with each other. Reproducible twofold increases and decreases were observed in a few genes. We judged these alterations to be artificial errors and, as such, systematic biases. However, when KK46 and KK335 (*dam-16::kam*) were compared, reproducible twofold increases and decreases were observed in more than 90 genes. These changes are considerably more frequent than those resulting from systematic biases. On the basis of these observations, we recognized spots to represent a significant alteration in transcription in the mutant compared with the wild type when the following conditions were fulfilled for the four independent data obtained for each gene: for group 1 spots, if the KK335 Cy5/KK46 Cy3 ratio was reproducibly less than 0.5, indicating a negative fold difference, or reproducibly more than 2.0, indicating a positive fold difference; for group 2 spots (a) if the Cy3 signal exceeded the high-intensity cut-off (>1000), whereas the Cy5 signal showed no significant intensity, indicating a higher transcript level in the wild type, or (b) if the Cy5 signal exceeded the high-intensity cut-off, whereas the Cy3 signal showed no significant intensity, indicating a higher transcript level in the mutant. This method was used to judge the expression alteration of all genes included in groups 1 and 2. (ii) The second method used to locate significant differences in the transcript levels of each gene in the mutant was by more statistical methods. This could be done only with the group 1 genes. We calculated the consistency of the differential expression in the four data

obtained per gene (two separate RNA preparations obtained from each strain were tested twice, thus yielding four independent spots per gene) using Wilcoxon signed rank test on the PYTHON program. Thus, spots with a significantly ( $P < 0.01$ ) lower ( $<0.5$ , i.e. a negative fold difference) or higher ( $>2$ , i.e. a positive fold difference) KK335 Cy5/KK46 Cy3 ratio were considered to be real differences.

To estimate the reproducibility of the independently performed experiments, we calculated Pearson's correlation coefficient for the two experiments performed with each of the two mRNA preparations by assessing the total gene expression profiles obtained with the same growth phase–aerobicity combination (i.e. aerobic log, aerobic stationary, low aerobic log and low aerobic stationary) (these data can be found at [http://ecoli.aist-nara.ac.jp/xp\\_analysis/dam/overlaps.html](http://ecoli.aist-nara.ac.jp/xp_analysis/dam/overlaps.html)). If the expression profiles between two independent experiments per experimental condition are similar, one would expect a high correlation coefficient. Indeed, a high value was obtained for three of the four growth conditions (aerobic log, 0.45; aerobic stationary, 0.55; low aerobic log, 0.70). In addition, the value was moderate but still significant for low aerobic stationary conditions (0.24). Thus, the microarray data in this analysis appear to be reproducible.

For each growth condition, the number of genes that could be included in groups 1 and 2 after one hybridization exceeded 2400 ( $>$  about 60% of the total spotted genes on the slide). When the four individual hybridizations were examined, transcript profiles of more than 90 genes (that is at least 2% of total genes) were found to be reproducibly and significantly altered in the mutant strain. These data are available on our web site at [http://ecoli.aist-nara.ac.jp/xp\\_analysis/dam/](http://ecoli.aist-nara.ac.jp/xp_analysis/dam/).

#### Northern blot analysis

The analysis was performed using specific probes essentially as described previously (Ito *et al.*, 1993). Total RNA was prepared as described above and separated using 1.0% agarose gel electrophoresis with formamide and transferred to a Hybond-N<sup>+</sup> membrane (Amersham Pharmacia). All DNA fragments used for this analysis were amplified by PCR using ORF-specific primers (see <http://ecoli.aist-nara.ac.jp/PRIMER/index.html>). The DNA probes were labelled with [<sup>32</sup>P]-dCTP (Amersham Pharmacia) by a random priming kit (Takara). The blotted membrane was hybridized with each labelled probe, washed, dried and visualized with a Fuji bioimaging analyser (BAS2000, Fuji).

#### Motility analysis

Overnight cultures of cells were spotted onto semi-solid agar plates containing 0.8% nutrient broth (Difco), 8% gelatin (Wako, Japan) and 0.4% agarose (Difco). The plates were incubated at 30°C for 6 h and photographed.

#### Proteome analysis by RFHR 2-D PAGE

*Escherichia coli* were grown under the same aerobic condi-

tions used for the microarray analysis. An aliquot of 400 ml of cells was collected in the cold at log phase and at stationary phase grown under aerobic and low aerobic conditions. Cell growth was stopped by ice and azide (final 0.01 M). About 1.0–1.2 g of wet cells was suspended in 1.5 ml of buffer 1 [100 mM CH<sub>3</sub>COONH<sub>4</sub>, 15 mM (CH<sub>3</sub>COO)<sub>2</sub>Mg, 20 mM Tris-HCl, pH 7.6, 6 mM β-mercaptoethanol, and 0.5 mM phenylmethylsulphonyl fluoride (PMSF)], sonicated for 15 min at 20 kHz with 50% duty cycle (on for 7.5 s then off for 7.5 s) (Bioruptor UCD-200TM, Cosmo Bio) in the cold and then centrifuged for 10 min at 900 *g*. The supernatant (sup) was centrifuged for a further 20 min at 10 000 *g*. The pellet (ppt) arising from this latter centrifugation was resuspended in 2 ml of buffer 1, centrifuged for 20 min at 10 000 *g* and then homogenized in 0.5–0.7 ml of buffer 1. This fraction was defined as 'CD'. The sup after centrifugation for 20 min at 10 000 *g* was centrifuged further for 180 min at 100 000 *g*. The sup was defined as 'PRS', and the ppt was resuspended in 0.5 ml of buffer 1 and centrifuged for 10 min at 17 000 *g*. The sup was defined as 'CR'.

The PRS, CD and CR fractions were resuspended in a solution of 67% acetic acid and 33 mM MgCl<sub>2</sub> and centrifuged for 10 min at 10 000 *g*. The ppts were resuspended in the same buffer, and the elution procedure was repeated. The two sups obtained after this were combined and desalted by Sephadex G-25 (Medium). The samples were then lyophilized. Lyophilized protein ( $\approx$  1–2 mg per gel) was analysed by RFHR 2-D PAGE essentially as described previously (Wada, 1986; see <http://www.osaka-med.ac.jp/~yhide/index.htm>), except that the volume of glacial acetic acid used in the sample charging buffer (50 $\times$ ) was 7.4 ml, not 74 ml, and a gel thickness of 2 mm was used to improve the resolution.

After RFHR 2-D PAGE, the gels were stained with CBB (Coomassie brilliant blue R250) to visualize the proteins. The protein spots were measured by Personal Densitometer S1 (Molecular Dynamic Japan), quantified using IMAGE QUANT (Molecular Dynamic Japan), and the density of the protein spots in the wild-type strain was compared with that of the *dam* strain. Protein spots that were increased in the *dam* strain compared with the wild-type strain were identified by the gene–protein index for RFHR 2-D PAGE (<http://ecoli.aist-nara.ac.jp/docs/GB4/search/proteome/index.html>). The spots not included in the gene–protein index were identified by peptide sequencing or matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). Briefly, this meant that, after the RFHR 2-D PAGE, either proteins were blotted to a polyvinylidene difluoride (PVDF) membrane for determination of N-terminus amino acid sequences using a Shimadzu PPSQ-23 peptide sequencer or the protein spots were digested in proteinase (Endoproteinase LysC) for MALDI-TOF MS analyses (Voyager DETMPRO, Applied Biosystems).

The density of each protein was normalized by the control proteins, molecular weight markers (molecular weight 2512–16949, Amersham Pharmacia Biotech) and the internal proteins, which showed very little difference in density between the wild-type strain and the *dam* strain. We used two or three proteins as control proteins in each gel but obtained equivalent results. We isolated two to four independent protein preparations and tested the protein prepa-

rations from each growing condition three to six times on 2-D PAGE. Reproducibly and significantly altered spots were selected. The density of each protein in the mutant was related to the wild-type protein density by expressing the two values as a ratio (KK335/KK46). Table 2 indicates selected proteins whose ratio of the relative density of protein from KK335 (*dam-16::kam*) relative to that of the wild-type strain exceeds 1.3.

#### Software analysis

The *Kyoto Encyclopedia of Genes and Genomes* (KEGG) pathway database (Kanehisa and Goto, 2000) was used to analyse the Dam-regulated biological cascade. GENESPRING version 3.0 (Silicon Genetics) was used to locate CRP and Fnr recognition sequences, to discover where GATC sequences overlap with these recognition sequences and to perform statistical analysis. Statistical analyses were also performed by EXCEL 2000 and the PYTHON program.

#### Acknowledgements

We thank M. Kitagawa, T. Nakamichi-loka, E. Inamoto, H. Toyonaga and S. Kanata for the *E. coli* clone bank essential for the construction of DNA microarray, E. Boye for the *dam* mutant strain, G. Kobayashi and A. Wada for proteome analysis, M. Ueta for laboratory supplies, M. Kitagawa and T. Horiuchi for helpful discussions, and T. Yura for warm support and critical reading of the manuscript. This work was supported by the CREST programme of Japan Science and Technology and Grants-in-Aid for Scientific Research on Priority Areas, 'Genome Science' from the Ministry of Education, Science, Sports and Culture of Japan.

#### References

- Aiba, H. (1985) Transcription of the *Escherichia coli* adenylate cyclase gene is negatively regulated by cAMP receptor protein. *J Biol Chem* **260**: 3063–3070.
- Arfin, S.M., Long, A.D., Ito, E.T., Toller, L., Riehle, M.M., Paegle, E.S., and Hatfield, G.W. (2000) Global gene expression profiling in *Escherichia coli* K-12: the effects of integration host factor. *J Biol Chem* **275**: 29672–29684.
- Basby, S., and Kolb, A. (1996) The CAP modulon. In *Regulation of Gene Expression in Escherichia coli*. Lin, E.C.C., and Lynch, A.S. (eds). Texas: R.G. Landes, pp. 255–277.
- Biville, F., Laurent-Winter, C., and Danchin, A. (1996) *In vivo* positive effects of exogenous pyrophosphate on *Escherichia coli* cell growth and stationary phase survival. *Res Microbiol* **147**: 597–608.
- Blattner, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., et al. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453–1474.
- 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.
- Braaten, B.A., Nou, X., Kaltenbach, L.S., and Low, D.A. (1994) Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli*. *Cell* **76**: 577–588.
- Choe, L.H., Chen, W., and Lee, K.H. (1999) Proteome analysis of factor for inversion stimulation (Fis) overproduction in *Escherichia coli*. *Electrophoresis* **20**: 798–805.
- De Reuse, H., and Danchin, A. (1991) Positive regulation of the *pts* operon of *Escherichia coli*: genetic evidence for a signal transduction mechanism. *J Bacteriol* **173**: 727–733.
- DeRisi, J.L., Iyer, V.R., and Brown, P.O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**: 680–686.
- 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 USA* **96**: 11578–11583.
- Hale, W.B., van der Woude, M.W., and Low, D.A. (1994) Analysis of nonmethylated GATC sequence sites in the *Escherichia coli* chromosome and identification of sites that are differentially methylated in response to environmental stimuli. *J Bacteriol* **176**: 3438–3441.
- 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.
- Henaut, A., Rouxel, T., Gleizes, A., Moszer, I., and Danchin, A. (1996) Uneven distribution of GATC sequence motifs in the *Escherichia coli* chromosome, its plasmids and its phages. *J Mol Biol* **257**: 574–585.
- Hiraga, S., Ichinose, C., Onogi, T., Niki, H., and Yamazoe, M. (2000) Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of *oriC* copies in *Escherichia coli*. *Genes Cells* **5**: 327–341.
- Ito, K., Kawakami, K., and Nakamura, Y. (1993) Multiple control of *Escherichia coli* lysyl-tRNA synthetase expression involves a transcriptional repressor and a translational enhancer element. *Proc Natl Acad Sci USA* **90**: 302–306.
- Jones, B.D., Lee, C.A., and Falkow, S. (1992) Invasion by *Salmonella typhimurium* is affected by the direction of flagellar rotation. *Infect Immun* **60**: 2475–2480.
- Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27–30.
- Laurent-Winter, C., Ngo, S., Danchin, A., and Bertin, P. (1997) Role of *Escherichia coli* histone-like nucleoid-structuring protein in bacterial metabolism and stress response-identification of targets by two-dimensional electrophoresis. *Eur J Biochem* **244**: 767–773.
- Li, J., Kustu, S.T., and Stewart, V. (1994) *In vitro* interaction of nitrate-responsive regulatory protein NarL with DNA Target sequences in the *fdnG*, *narG*, *narK* and *frdA* operon control regions of *Escherichia coli* K-12. *J Mol Biol* **241**: 150–165.
- Lin, R.J., Capage, M., and Hill, C.W. (1984) A repetitive DNA sequence, *rhs*, responsible for duplications within the *Escherichia coli* K-12 chromosome. *J Mol Biol* **177**: 1–18.
- Liu, S.L., Ezaki, T., Miura, H., Matsui, K., and Yabuuchi, E. (1988) Intact motility as a *Salmonella typhi* invasion related factor. *Infect Immun* **56**: 1967–1973.
- Macnab, R.M. (1996) Flagella and motility. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Neidhardt,

- F., *et al.* (eds). Washington, DC: American Society for Microbiology Press, pp. 123–145.
- Marinus, M.G. (1996) Methylation of DNA. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Neidhardt, F., *et al.* (eds). Washington, DC: American Society for Microbiology, pp. 782–791.
- Marinus, M.G. (2000) Recombination is essential for viability of an *Escherichia coli dam* (DNA adenine methyltransferase) mutant. *J Bacteriol* **182**: 463–468.
- Mori, H., Isono, K., Horiuchi, T., and Miki, T. (2000) Functional genomics of *Escherichia coli* in Japan. *Res Microbiol* **151**: 121–128.
- Nou, X., Braaten, B., Kaltenbach, L., and Low, D.A. (1995) Differential binding of Lrp to two sets of pap DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*. *EMBO J* **14**: 5785–5797.
- Nyström, T. (1995) Glucose starvation stimulon of *Escherichia coli*: role of integration host factor in starvation survival and growth phase-dependent protein synthesis. *J Bacteriol* **177**: 5707–5710.
- Nyström, T., Larsson, C., and Gustafsson, L. (1996) Bacterial defense against aging: role of the *Escherichia coli* ArcA regulator in gene expression, readjusted energy flux and survival during stasis. *EMBO J* **15**: 3219–3228.
- Oh, T.J., and Kim, I.G. (1999) Identification of genetic factors altering the SOS induction of DNA damage-inducible yebG gene in *Escherichia coli*. *FEMS Microbiol Lett* **177**: 271–277.
- Parker, B., and Marinus, M.G. (1988) A simple and rapid method to obtain substitution mutation in *Escherichia coli*: isolation of a *dam* deletion/insertion mutation. *Gene* **73**: 531–535.
- Peterson, K.R., Wertman, K.F., Mount, D.W., and Marinus, M.G. (1985) Viability of *Escherichia coli* K 12 DNA adenine methylase (*dam*) mutants requires increased expression of specific genes in the SOS regulon. *Mol Gen Genet* **201**: 14–19.
- Plumbridge, J., and Kolb, A. (1991) CAP and Nag repressor binding to the regulatory regions of the *nagE-B* and *manX* genes of *Escherichia coli*. *J Mol Biol* **217**: 661–679.
- Resnik, E., Pan, B., Ramani, N., Freundlich, M., and LaPorte, D.C. (1996) Integration host factor amplifies the induction of the *aceBAK* operon of *Escherichia coli* by relieving IclR repression. *J Bacteriol* **178**: 2715–2717.
- Richmond, C.S., Glasner, J.D., Mau, R., Jin, H., and Blattner, F.R. (1999) Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res* **27**: 3821–3835.
- Riley, M., and Labeledan, B. (1996) *Escherichia coli* gene products: physiological functions and common ancestries. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Neidhardt, F.C., Curtiss, R., Gross, C., Ingraham, J.L., Lin, E.C.C., Low, K.B., *et al.* (eds). Washington, DC: American Society for Microbiology Press, pp. 2118–2202.
- Tao, H., Bausch, C., Richmond, C., Blattner, F.R., and Conway, T. (1999) Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media. *J Bacteriol* **181**: 6425–6440.
- Tavazoie, S., and Church, G.M. (1998) Quantitative whole genome analysis of DNA protein interactions by *in vivo* methylase protection in *E. coli*. *Nature Biotechnol* **16**: 566–571.
- Taylor, B.L., and Zhulin, I.B. (1998) In search of higher energy: metabolism-dependent behaviour in bacteria. *Mol Microbiol* **8**: 683–690.
- Taylor, B.L., Miller, J.B.O., Warrick, H.M., and Koshland, D.E., Jr (1979) Electron acceptor taxis and blue light effect on bacterial chemotaxis. *J Bacteriol* **140**: 567–573.
- Taylor, B.L., Zhulin, I.B., and Johnson, M.S. (1999) Aerotaxis and other energy-sensing behavior in bacteria. *Annu Rev Microbiol* **53**: 103–128.
- Van Bogelen, R.A., and Neidhardt, F. (1990) Ribosomes as sensors of heat and cold shock in *Escherichia coli*. *Proc Natl Acad Sci USA* **87**: 5589–5593.
- Wada, A. (1986) Analysis of *Escherichia coli* ribosomal proteins by an improved two dimensional gel electrophoresis. I. Detection of four new proteins. *J Biochem* **100**: 1583–1594.
- Wang, M.X., and Church, G.M. (1992) A whole genome approach to *in vivo* DNA protein interactions in *E. coli*. *Nature* **360**: 606–610.
- van der Woude, M., Hale, W.B., and Low, D.A. (1998) Formation of DNA methylation patterns: nonmethylated GATC sequences in *gut* and *pap* operons. *J Bacteriol* **180**: 5913–5920.
- Yamanaka, K., Ogura, T., Niki, H., and Hiraga, S. (1996) Identification of two new genes, *mukE* and *mukF*, involved in chromosome partitioning in *Escherichia coli*. *Mol Gen Genet* **250**: 241–251.
- Zheng, M., Wang, X., Templeton, L.J., Smulski, D.R., LaRossa, R.A., and Storz, G. (2001) DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J Bacteriol* **183**: 4562–4570.