Nucleic Acids Research, 2001, Vol. 29, No. 24 **5195–5206**

Comparative analysis of FUR regulons in gamma-proteobacteria

E. M. Panina*, A. A. Mironov[1] and M. S. Gelfand[1]

State Scientific Center GosNIIGenetika and [1]IntegratedGenomics-Moscow, 1st Dorozhny prospect, Moscow 113545, Russia

Received June 20, 2001; Revised and Accepted October 16, 2001

ABSTRACT

Iron is an essential element for the survival and pathogenesis of bacteria. The strict control of iron homeostasis is mediated by the FUR repressor, which is highly conserved among various bacterial species. Here we apply the comparative genomics approach to analyze candidate Fur-binding sites in the genomes of *Escherichia coli* (K12 and O157:H7), *Salmonella typhi*, *Yersinia pestis* and *Vibrio cholerae*. We describe a number of new loci encoding siderophore biosynthesis and transport proteins. A new regulator of iron-acquisition systems was found in *S.typhi*. We predict FUR regulation for several virulence systems. We also predict FUR regulation for the chemotaxis system of *V.cholerae* that is probably involved in the process of pathogenesis.

INTRODUCTION

Iron is necessary for the growth of most bacteria. Therefore, efficient iron-acquisition mechanisms must be employed when bacteria are facing iron-restricted conditions such as the host environments for pathogenic strains. Many microorganisms secrete siderophores, high-affinity iron III chelators. Then the Fe(III)–siderophore complexes are taken up into the cell via specific transport systems usually including a TonB-dependent outer membrane receptor and an ATP-binding type transporter (1).

Escherichia coli is known to synthesize two types of siderophores: enterobactin and aerobactin, and it also utilizes such exogenous siderophores as ferrichrome, ferridicitrate, coprogen and rhodotorulic acid (1). In the *E.coli* genome, the iron uptake systems are encoded by more than 20 genes that form several operons: the aerobactin biosynthesis and uptake system *iucABCDiutA* (2); the ferrichrome-uptake system *fhuACDB* (3); the uptake/sensory system for the ferridicitrate complexes, *fecIRABCDE* (4); and large locus-containing genes for enterobactin biosynthesis (*entABCDEFG*) and uptake (*fepABCDEG*) organized in five small operons (1). There are also several other proteins involved in iron acquisition: FhuE is a coprogen and rhodotorulic acid receptor, and its gene *fhuE* is located separately from other *fhu* genes (5); Fes, encoded by the first gene of the *fes-entF-fepE* operon, is a protein that hydrolyzes enterobactin, thus facilitating Fe(III) reduction to Fe(II) (6). Finally, the TonB–ExbBD complex provides energy that allows the ferrisiderophores to enter the periplasm through the outer membrane (7).

Although iron is indispensable, its excess in the cytoplasm is toxic for the cell, as it catalyzes the Fenton reaction, leading to formation of hydroxyl radicals (OH·) (8). Thus, iron homeostasis is strictly controlled. In Enterobacteriaceae this control is mediated by Fur (Fe uptake regulator) and its orthologs (2,9). In *E.coli*, in the presence of iron, Fur binds DNA and represses the transcription of all genes involved in siderophore biosynthesis and transport except for *tonB* and *exbBD* (2). A 19-bp consensus Fur-binding site (FUR-box) GATAATGATAATCATTATC was proposed to be based on DNase I protection and footprinting experiments (10,11).

The availability of complete genomes of related bacteria allows one to apply the comparative approach to the analysis of regulatory patterns. An important advantage of simultaneous analysis of several genomes is the possibility of distinguishing true sites, occurring upstream of orthologous genes, from false positives scattered at random across the genome (12,13). The latter technique allowed us and other groups to analyze several regulons and to obtain a number of promising predictions (14–18). Here we apply this technique to analyze Fur regulation in the genomes of *E.coli* (K12 and O157:H7), *Salmonella typhi*, *Yersinia pestis* and *Vibrio cholerae*. As iron-acquisition systems are often located on plasmids or within transposable pathogenicity islands, they are a frequent subject of horizontal transfer (19,20). Thus, these systems are widely distributed among various bacterial species, but even closely related genomes can contain different sets of systems. This makes the standard orthology analysis inapplicable in many cases.

On the other hand, the Fur protein is conserved not only within Gram-negative, but also within Gram-positive bacteria (21,22). Its recognition signal is conserved as well (23,24).

Thus, in this study we modified the basic comparative technique (13), accepting a site upstream of a functionally relevant gene if either of the following conditions was met: (i) there were candidate sites upstream of all orthologous genes in genomes from the same taxonomic group, gamma-proteobacteria; (ii) there was an experimentally verified site upstream of a closely homologous gene from another genome; (iii) in case there was an orthologous gene with no candidate site in the same taxonomic group, there were at least two candidate sites upstream of homologous genes from distant species and (iv) in the case of no orthologous genes being found, the

*To whom correspondence should be addressed. Tel: +7 095 315 01 56; Fax: +7 095 315 05 01; Email: katya@ekpanina.mccme.ru
Correspondence may also be addressed to M. S. Gelfand. Tel: +7 095 315 0156; Fax: +7 095 315 0501; Email: misha@imb.imb.ac.ru

Table 1. The decision procedure

Close genomes from the same taxonomic group	Experimentally verified sites upstream of some closely homologous genes	Distant genomes		
		Candidate sites upstream of some homologous genes	No candidate sites upstream of homologous genes	No homologous genes
Candidate sites upstream of *all* orthologous genes	Accept a candidate site	Accept a candidate site	Accept a candidate site	Accept a candidate site
Candidate sites upstream of *some* orthologous genes	Accept a candidate site	Accept a candidate site	Reject a candidate site	Reject a candidate site
No orthologous genes, but relevant function	Accept a candidate site	Accept a candidate site	Reject a candidate site	Accept a candidate site

Table 2. The training set

Gene	Site
iucA	GATAATGAGAATCATTATT
Cir	GATAATTGTTATCGTTTGC
Fur	TATAATGATACGCATTATC
FecI	TGTAATGATAACCATTCTC
sltA	GAATATGATTATCATTTTC
fepB	GAAAATGAGAAGCATTATT
entC	ATAAATGATAATCATTATT
sodA	GATAATCATTTTCAATATC

function was directly involved in iron acquisition (Table 1). Additionally, a chain of functionally related genes transcribed in the same direction and with small intergenic distances was assumed to form a candidate operon, and a candidate site upstream of this operon was accepted as the one regulating all genes in the chain. Thus, we identified a number of potentially Fur-regulated genes encoding ferrisiderophore transport systems and other proteins involved in the iron acquisition.

Many bacterial pathogens use the low concentration of iron present in the host as an environmental clue to enhance the expression of a wide variety of bacterial toxins and other virulence determinants (25). In various bacterial species, chemotaxis plays an important role during pathogenesis. In particular, the *E.coli* chemotaxis system includes methyl-accepting chemotaxis proteins (MCPs), *che*-system for signal transduction and MotAB, FliMNG system for flagellar rotation (see Fig. 3A) (26). Here we present a new Fur-dependent chemotaxis system of this type, which is probably used by *V.cholerae* during its pathogenesis. Some other virulence systems are also proposed to be under the Fur control.

MATERIALS AND METHODS

The complete genome sequences of *E.coli* (both K12 and O157:H7), *Haemophilus influnzae*, *V.cholerae* and *Pseudomonas aeruginosa* were extracted from GenBank (27). Partially sequenced genomes of *S.typhi*, *Y.pestis* and *Shewanella putrefaciens* were extracted from the TIGR WWW site (http://www.tigr.org). In several cases, additional analyses were done on sequence fragments extracted from GenBank (27).

Known *E.coli* Fur-binding sites were taken (1) and a signal recognition profile was constructed. The quality of the profile was defined as its informational content (28). Then we retained all the sites from the first sample with scores exceeding 4.5 (Table 2) and computed the profile used in this study.

Positional nucleotide weights in this profile are defined as:
$W(b,k) = \log[N(b,k) + 0.5] - 0.25\Sigma_{i=A,C,G,T}\log[N(i,k) + 0.5]$,
where $N(b,k)$ is the count of nucleotide b at position k. The score of a candidate site is calculated as the sum of the respective positional nucleotide weights:
$Z(b_1 \ldots b_L) = \Sigma_{k=1 \ldots L} W(b_k,k)$,
where L is the length of the site. The cut-off value for candidate FUR-boxes in *E.coli*, *S.typhi* and *Y.pestis* was 4.5. The cut-off value for candidate FUR-boxes in *V.cholerae* was 4.25.

The DNA-binding motif in the FUR protein has been identified recently (29). This motif is strictly conserved in the Fur proteins from *E.coli*, *S.typhi*, *Y.pestis* and *V.cholerae* (Fig. 1). This allows us to use the same profile for the recognition of FUR-boxes in all studied genomes. However, several amino acids close to the recognition helix in *V.cholerae* Fur protein are different from those in the other studied genomes. As the DNA recognition domain in the Fur protein has a non-classical, not absolutely understood structure (29), these amino acids might play some role in the sequence-specific DNA binding. Thus, we decreased the cut-off for the *V.cholerae* genome to 4.25.

The cut-off 4.5 is passed by 12 out of 16 experimentally confirmed Fur-binding sites from the *E.coli* genome. Decreasing the cut-off to 4.25 in *E.coli* doubles the number of genes with candidate FUR-boxes. However, few genes with the site score <4.5 have a function related to iron metabolism. Nevertheless, we retained genes with sites scoring <4.5 (but >4.00) if they had orthologs with strong candidate sites, or were known to be regulated by Fur. We also recorded low-scoring candidate Fur-boxes within large iron-acquisition loci.

At the cut-off 4.5 we found four FUR-boxes out of 60 predicted (after comparative analysis) lie within coding regions. Database similarity searches were done using BLAST (http://www.ncbi.nlm.nih.gov/blast/). Multiple protein alignments and phylogenetic trees were constructed using CLUSTALX (30). Genomic analyses (protein similarity search, analysis of orthology, DNA profile search) were done using GenomeExplorer (31). The gene names suggested in this paper, based on homology and regulatory analyses, are marked by stars.

RESULTS

The profile made using known *E.coli* Fur-binding sites (see Materials and Methods) was applied to detect new candidate FUR-boxes in the genomes of *E.coli*, *S.typhi*, *Y.pestis* and *V.cholerae.*

Figure 1. The alignment of Fur proteins from *E.coli*, *S.typhi*, *Y.pestis* and *V.cholerae*. Amino acids of the DNA recognition helix are shown in yellow. Lys76, which also might interact directly with DNA, is shown in green.

The profile search in the *E.coli* K12 genome identified 73 genes having candidate FUR-boxes in the region (–300 … +20) relative to the start of translation. A complete table of orthologous genes was constructed for the genomes of *S.typhi*, *Y.pestis*, *V.cholerae*, *H.influenzae*, *P.aeruginosa* and *S.putrefaciens.* For those genes that had no orthologs or close homologs in the above genomes, homologs in distant species were identified using BLAST (http://www.ncbi.nlm.nih.gov/blast/). Then the *E.coli* K12 profile was used to identify the candidate FUR-boxes upstream of orthologous and homologous genes in the other genomes. As a result, six new candidate Fur-regulated genes were found.

The same procedure was applied to the genomes of *E.coli* O157:H7 (–300 … +50), *S.typhi* (–400 … +100 relative to the start codons), *Y.pestis* and *V.cholerae* (–500 … +100 relative to the start codons). The candidate sites within coding regions of upstream genes (deeper than 50 nt) were excluded. The number of genes with candidate FUR-boxes in the four genomes was 89, 75, 88 and 94, respectively.

In all five genomes, the profile search selected a number of experimentally confirmed FUR-boxes. They are not discussed below unless these genes have been used for the verification of candidate sites.

Below we carefully describe the predicted Fur-regulons from the genomic perspective. The functional overviews are given in the Discussion. The evolutionary perspective is presented in the Conclusion.

Candidate FUR-boxes in the *E.coli* K12 genome

Table 3 shows *E.coli* genes with new candidate FUR-boxes: *fptAybiX*, *ydiE*, *b1995*, *b3070*, *yhhX*.

The genes *fptAybiX* are likely to form an operon, as the length of the intergenetic region is only 6 nt. We have found a strong candidate FUR-box upstream of *fptA* (Table 3 and Fig. 2A). In the study by Oshima *et al.* (32), *fptA* was annotated as a Fe(III)–pyochelin receptor due to its similarity to the *P.aeruginosa* gene *piuA* with the same function. The corresponding locus in the *P.aeruginosa* genome is shown in Figure 2B. It also includes *piuC*, the ortholog of *ybiX* that encodes an iron-uptake factor, and a strong FUR-box (Table 3). The arrangement of genes in *P.aeruginosa* is different, as *piuA* and *piuC* are divergently transcribed (Fig. 2B). This system was shown to be regulated by the *P.aeruginosa* analog of the Fur (33). A locus similar to that of *P.aeruginosa* was found in the genome of *V.cholerae* (Fig. 2C). It also has a strong FUR-box (Table 3), although nothing is known about its regulation by Fur protein.

The *ydiE* and *b1995* genes are located separately in the genome of *E.coli*. Each of them has a strong upstream FUR-box (Table 3). BLAST revealed their similarity to the *Yersinia enterocolitica* haemin uptake proteins HemP (haemin-uptake factor) and HemR (haemin receptor), respectively (with *E*-values 9e-04 and 3e-07). In the genome of *Y.enterocolitica*, *hemR* and *hemP* form an operon. A FUR-box was described in its operator region and the iron-dependent transcription of *hemP* was shown (34). In the genome of *S.typhi* there is a locus similar to that described in *Y.enterocolitica*. It contains strong homologs of *hemP* and *hemR*, and a strong FUR-box (Table 3).

Table 3. The predicted FUR-boxes upstream of the iron-acquisition genes of *E.coli* and the FUR-boxes upstream of the homologous genes in other genomes

Gene	Organism	Score	Position	Site
fptA	EC	5.08	-128	GATAATGcTTATCAaaATt
piuA	PA	4.12**	-143	GccAATGATAtTgATTtgC
VC0284	VC	4.60	-87	GAaAATtAcTATCATTtgt
ydiE	EC	4.83	-42	GATAATaAgAATCATTgTt
b1995	EC	5.00	-28	GAaAATaATTATCATTAca
hemP	YE	4.00**	5	atgAtTGATAATgcTTATC
*hemP**	ST	4.66	-73	cAcAATGATTATtATTATC
b3070	EC	4.76	-64	aAcAATcATTATCATTtgC
viuB	VC	4.66	-143	GtTAATGATAtaCATTcTC
*viuB**	ST	4.49	-33	aAcAATcATTATCATTtaa
yhhX	EC	4.87	-231	GAgAATGATTATtATTgTC
*yhhX**	ST	4.84	-374	GcgAATGAgAATgATTATt
yvaA	BS	4.54	-109	GATAtTaATTccCATTATa

Genome abbreviation: EC, *E.coli* K12; ST, *S.typhi*; YP, *Y.pestis*; VC, *V.cholerae*; PA, *P.aeruginosa*; YE, *Y.enterocolitica*; BS, *B.subtilis*.
*Genes that have no common name or have been identified previously as open reading frames only. These genes are named after orthologs from other genomes.
**Experimentally confirmed Fur-binding sites.

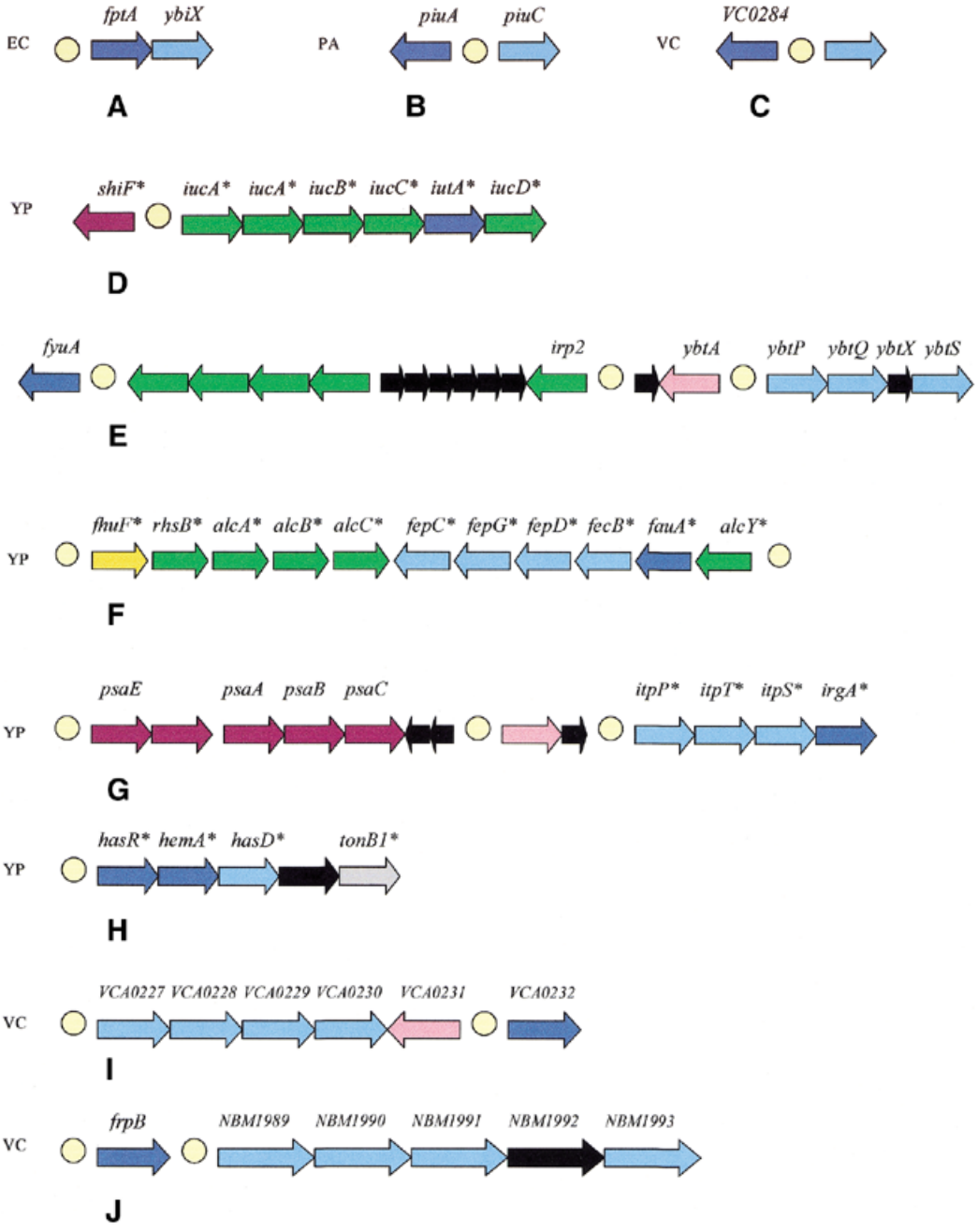

Figure 2. (A–C) The structure of three homologous iron-acquisition loci of: (**A**) *E.coli*, (**B**) *P.aeruginosae* and (**C**) *V.cholerae.* The gene *orf1* homologous to the iron uptake factors of *E.coli* and *P.aeruginosa* is missed in the annotation of the full genome of *V.cholerae*. (**D–F**) The structure of three siderophores-biosynthesis loci of *Y.pestis*. (**G**) The structure of *Y.pestis* locus that consists of the iron ABC transporter, the outer membrane receptor, the putative transcriptional regulator and the virulence-associated psa system. (**H**) The structure of the heme-acquisition locus of *Y.pestis*. (I and J) The structure of two iron-transport systems of (**I**) *V.cholerae* and (**J**) *N.meningitis*. Genes are shown by arrows, candidate FUR-boxes are shown by yellow circles. Blue arrow, the gene for iron ABC transporter; violet arrow, the gene for outer membrane ferrisiderophore receptor; lilac arrow, the gene for virulence-associated protein; pink arrow, the gene for transcriptional regulator; green arrow, the gene for siderophore-biosynthesis protein; yellow arrow, the gene for iron reductase; grey arrow, the gene for tonB protein; black arrow, other genes.

The next *E.coli* gene with a candidate FUR-box, *b3070* (Table 3), is orthologous to *viuB* from *V.cholerae* (*E*-value 4e-15). This gene encodes an iron-chelator utilization protein. In *V.cholerae*, *viuB* belongs to the iron-acquisition locus *vuiAiuBviuF* shown to be regulated by Fur (35). Transcription of *viuB* is iron dependent and it has a candidate FUR-box in the promoter region (Table 3) (36). The ortholog of *viuB* from *S.typhi* also has an upstream candidate FUR-box (Table 3).

The gene *yhhX* is a putative oxidoreductase. The possible role of oxidoreductases in the iron uptake is discussed below (see Discussion). We have found a candidate FUR-box

Table 4. The predicted FUR-boxes upstream of the iron-acquisition genes of *E.coli* O157:H7 and the FUR-boxes upstream of the homologous genes in other genomes

Gene	Organism	Score	Position	Site
Z1026	EC(pat)	5.08	-134	GATAATGcTTATCAaaATt
ydiE	EC(pat)	4.83	-36	GATAATaAgAATCATTgTt
Z3159	EC(pat)	5.00	-37	GAaAATaATTATCATTAca
yqjH	EC(pat)	4.76	-70	aAcAATcATTATCATTtgC
yhhX	EC(pat)	4.87	-238	GAgAATGATTATtATTgTC
iha	EC(pat)	4.92	-65	ataAATtATTcTCATTATt
*iha**	YP	4.78	7	GATAATtgaTATCATTtgC
Z2267	EC(pat)	5.20	-108	aATAATGATTAcCATTccC
*Z2267**	ST	5.27	-29	aATAATGATTATtATTcTC
Z2268	EC(pat)	5.20	-152	GggAATGgTAATCATTATt
*fyuA**	ST	5.27	-135	GAgATTaATAATCATTATt
fyuA	YP	4.69	-72	AgaAATagTTATCATTtTC
Z4385	EC(pat)	4.51	-52	GAgAATGAgTtTtATTtTt
yfmC	BS	5.62	-42	GATAATGATTcTCATTAct
VC0776	VC	4.49	1	cgaAATGAgAATCAaTATt
chuT	EC(pat)	5.12	-37	GATAATcATgATCATTcTC
chuA	EC(pat)	5.55	-313	GATAATaATTcTCATTATA

Genome abbreviation: EC(pat), *E.coli* O157:H7; see also Table 3.
*Genes that have no common name or have been identified previously as open reading frames only. These genes are named after orthologs from other genomes.

upstream of this gene and upstream of its homolog *yvaA* in the genome of *Bacillus subtilis* (Table 3). The *E.coli* genome contains a gene *b1624* with stronger similarity to *yvaA* (44% identity compared with 30%) and no FUR-box. The *S.typhi* ortholog of *yhhX* also has a candidate FUR-box (Table 3). After this study was completed, the FUR-box upstream of *yhhX* was experimentally confirmed (37).

Candidate FUR-boxes in the *E.coli* O157:H7 genome

Table 4 shows *E.coli* O157:H7 genes with new candidate FUR-boxes: *Z2267*, *Z2268*, *chuTWXYUV**, *z4912chuAS*, *Z1026ybiX*, *Z3159*, *Z1178*, *Z1617*, *ydiE*, *yqjH*, *Z4385-Z4382*, *Z4386*, *yhhX*.

Some of these genes have already been described in the above paragraph. *Z1026* and *ybiX* from *E.coli* O157:H7 are orthologs of the *fptA* and *ybiX* from *E.coli* K12, respectively; *Z3159* and *ydiE* are orthologs of *b1995* and *ydiE*; *yqjH* is orthologous to *b3070*; *yhhX* has the same name in both genomes.

A new gene *iha* encoding an adhesin was recently discovered in the pathogenic strain of *E.coli* O157:H7 (38). Actually, in the genome of *E.coli* O157:H7 there are two identical genes for the Iha adhesin, *Z1178* and *Z1617*. We have found strong FUR-boxes upstream of both genes (Table 4), and the FUR-box is conserved upstream of the orthologous gene in the genome of *Y.pestis* (see below).

Z2267 and *Z2268* form a divergon. *Z2267* encodes a putative receptor with a candidate role in enterotoxin production, determined by homology to the *S.typhi* gene with the same function. *Z2268* encodes a putative receptor for iron transport (as annotated in GenBank). A candidate FUR-box was observed in the intergenetic region of *Z2267* and *Z2268* (Table 4). It is conserved in the genomes of *S.typhi* and *Y.pestis*. In the latter case there are no orthologs for *Z2267* and the signal was found only upstream of *fyuA*, a *Z2268* ortholog with 24% identity.

The genes *Z4385–Z4382* encode an iron ABC transporter, whereas the divergent gene *Z4386* encodes an iron-compound receptor. We have found a candidate FUR-box in the *Z4385–Z4386* intergenic region (Table 4). The *Z4385–Z4382* operon is similar to the iron ABC transporter of *B.subtilis yfmCDEF* and to the ferric vibriobactin ABC transporter of *V.cholerae VC0776–VC0779* (although the arrangement of genes in both cases is different). Candidate FUR-boxes precede both *yfmC* and *VC0776*.

Escherichia coli O157:H7 also has a large locus of genes involved in iron acquisition with no orthologs in *E.coli* K12 and *S.typhi*. This locus includes: *chuTWXYUV** and divergent *Z4912-chuAS*. The first gene, *chuT*, encodes a putative periplasmic-binding protein, *chuW* encodes a putative oxygen independent coproporphyrinogen III oxidase and *chuX* and *chuY* are unknown open reading frames. Further, *chuU* encodes a putative permease of iron-compound ABC transport system, and *chuV** (*Z4919*) encodes a hypothetical protein highly similar to ATPases of iron ABC transporters. Finally, *Z4912* is an unknown open reading frame, *chuA* encodes an outer membrane heme/hemoglobin receptor, and *chuS* encodes a putative heme/hemoglobin transport protein. A strong candidate FUR-box was found in the *Z4912-chuT* intergenic region. This system was also found in the genomes of *Y.pestis* and *V.cholerae* and in a reduced form in the genomes of *S.putrefaciens* and *P.aeruginosa*. The arrangement of genes in each case is unique, but a candidate FUR-box is always present in the regulatory region. There is a large intergenic space of 216 nt between *Z4912* and *chuA*. We have found another strong candidate FUR-box upstream of *chuA*, although it overlaps with *Z4912*.

Candidate FUR-boxes in the *S.typhi* genome

The *S.typhi* genes with new candidate FUR-boxes are *feoAB*, *sodA*, *hmsT**, *foxR** (Table 5).

Two systems, *feoAB* for Fe(II) transport and *sodA* encoding the superoxide dismutase, are known to be regulated by Fur in *E.coli* (1). There are candidate FUR-boxes upstream of the *sodA* orthologs in the genomes of *S.typhi*, *Y.pestis* and *V. cholerae* (Table 5). Candidate FUR-boxes upstream of the orthologs of *feoAB* were found in the genomes of *S.typhi* and *Y.pestis* (Table 5).

The gene *hmsT** from *S.typhi* is an ortholog of *hmsT* from *Y.pestis* that encodes a haemin storage protein. Both genes have candidate FUR-boxes (Table 5). In *Y.pestis*, the *hms*-phenotype was shown to be regulated by Fur (39). In the genome of *S.typhi* this gene is followed by an ortholog of *fhuF*, the Fur-regulated gene from *E.coli*.

FoxR* of *S.typhi* is an AraC/XylS-type regulator. This gene has a candidate FUR-box. There are candidate FUR-boxes upstream of closely related regulators in other genomes (*alcR*, *ybtA*, *pchR*; Table 5). The gene *alcR* from *Bordetella pertussis* was shown to be regulated by Fur (40); *ybtA* from *Y.pestis* belongs to the yersiniabactin biosynthesis locus and and its expression is Fur dependent (41); *pchR* from *P.aeruginosa* is likely to be regulated by Fur (42). Besides, all three proteins

Table 5. The predicted FUR-boxes upstream of the iron-acquisition genes of *S.typhi* and the FUR-boxes upstream of the homologous genes in other genomes

Gene	Organism	Score	Position	Site
*hmsT**	ST	5.68	-221	GATAATGATAAcCAaTATC
hmsT	YP	5.58	-264	GATAATGATAATCATaAcC
foxR	ST	4.84	-267	GATAATtATTcTCATctTt
ybtA	YP	4.73**	-91	GtgAATaATAAcCATTATC
alcR	BP	4.56**	-28	GcgAATGAaTtgCATTATC
pchR	PA	4.63	~-120	GATAATaTaTcTCATTtcC
*sodA**	ST	5.25	-94	GATAATcATTnTCAaTATC
sodA	EC	5.16**	-94	GATAATcATTtTCAaTATC
*sodA**	VC	5.20	-25	GtTAATGATAtTaATTATC
*sodA**	YP	5.16	-36	GATAATcATTtTCAaTATC
feoA	EC	4.64**	-133	agaAAccATTcTCATTATC
*feoA**	ST	4.64	-194	agaAAccATTcTCATTATC
*feoA**	YP	4.76	-64	GtTAATaATTATtATTATC

Genome abbreviation: BP, *B.pertussis*; see also Table 3.
*Genes that have no common name or have been identified previously as open reading frames only. These genes are named after orthologs from other genomes.
**Experimentally confirmed Fur-binding sites.

regulate transcription of Fur-regulated iron-acquisition systems. AlcR and YbtA act in *cis*, the target of PchR is located not further than 14 kb from the *pchR* gene. Given these observations, we propose the *foxA* gene as a candidate target for the regulation by FoxR*. Indeed, it is located immediately downstream of *foxR*, it has a strong FUR-box (43) and the product of *foxA* functions as a ferrisiderophore receptor.

Candidate FUR-boxes in the *Y.pestis* genome

Table 6 shows *Y.pestis* genes with new candidate FUR-boxes: *itpPTS**, *itsPTUS**, *itrAS**, *irp2*, *iucABCDiutA**, *alcABC**, *aclY**, *psaE, iha**, *hemN**, *fhuE**, *omrA**, *hasR** and *yebN**.

We propose the name Itp* (iron transport) for a new candidate Fe(III) ABC transporter, which was found in the *Y.pestis* genome. The Itp system consists of a periplasmic-binding protein (ItpP*), a transmembrane protein (ItpT*) homologous to the members of the FecC/D iron permease family (BLAST *E*-value 3e-47) and an ATPase (ItpS*) homologous to the ATP-binding protein VC0779 of the *V.cholerae* ferric vibriobactin ABC transporter (BLAST *E*-value 3e-28). Besides, the gene located immediately downstream of *itpS** has a strong similarity to the iron-regulated outer membrane ferrisiderophore receptor *irgA* from *V.cholerae*. There is a candidate FUR-box upstream of *itpP** (Table 6).

One more candidate Fe(III) ABC transporter found in the genome of *Y.pestis* was named Its* (iron transport system). It includes a periplasmic protein (ItsP*), two transmembrane proteins (ItsT*, ItsU*) and an ATPase (ItsS*). There is a candidate FUR-box upstream of *itsP** (Table 6). Iron transporter systems similar to Its exist in the genomes of *V.cholerae* and *Vibrio angulliarrium* (*fatDCBA*). The gene *fatB* is the ortholog of *itsP*. The Fur-dependent regulation of *fatBA* in *V.angulliarrium* was shown in Actis *et al*. (44). Although Fur does not bind to

Table 6. The predicted FUR-boxes upstream of the iron-acquisition genes of *Y.pestis* and the FUR-boxes upstream of the homologous genes in other genomes

Gene	Organism	Score	Position	Site
*itpP**	YP	4.84	-31	GATAAgtATTATCATTtgC
*itsP**	YP	4.86	0	GgTAATaAaAATCATTAgC
*fatB**	VC	4.15	-14	ctTAATGAgAATaAgTATC
*itrA**	YP	4.78	-45	GATttTGATTATCATTcgC
*iucA**	YP	5.32	-64	GATAATGATAAcCAcTATt
iucA	EC	6.00**	-56	GATAATGAgAATCATTATt
irp2	YP	4.76	-64	GtTAATaATTATtATTATC
*alcY**	YP	4.59	-206	ataAATaATAATCATTtgC
alcA	BP	4.63**	-150	cATAATtATTcTCATTAat
psaE	YP	4.74	-219	tAaAATcAgTcTCATTtTa
*iha**	YP	4.78	7	GATAATtgaTATCATTtgC
iha	EC	4.92	-65	ataAATtATTcTCATTATt
*hemN**	YP	4.79	5	GcTtATGATTATCATTtaa
VCA0909	VC	4.30	-98	aATtATcAaTtgCATTATC
hemN	EC	no site		
*hemN**	ST	no site		
*hemN**	HI	no site		
*fhuF**	YP	4.94	-65	GcTAtTGATAATCAcTATC
fhuF	EC	5.68	-121	GATAATGATAAcCAaTATC
*omrA**	YP	5.34	-188	GAaAATGATAATCgTTtgC
*omrA**	PA	4.36	-165	GAgAtcGATAtgCATTtTC
*hasR**	YP	5.98	-341	GATtATGATAATCATTATC
hasR	PA	4.57**	-126	GcaAATGAaAActATTATC
*yyyN**	YP	6.02	-205	GATAATGAaAATCATTATC
yebN	EC	4.49	-259	ccaAATGAaAATCgTTATC
*ftn**	YP	5.00	-94	aATAATGtgAATgATTATC
ftn	EC	4.59	-638	aAaAATGATAATCAaaAaa
ftn	HI	4.76	-101	GtaAATGATAATaATTgTt
*ftn**	HP	3.91**	-375	ataAATcATAATCATtTtg

Genome abbreviations: HI, *H.influenzae*; HP, *Helicobacter pylori*; see also Tables 3 and 5.
*Genes that have no common name or have been identified previously as open reading frames only. These genes are named after orthologs from other genomes.
**Experimentally confirmed Fur-binding sites.

the *fatB* promoter, it regulates the expression of the latter indirectly by binding to the promoter of the antisense RNA. (45). We have found a candidate FUR-box upstream of the *itsP* ortholog in the genome of *V.cholerae* (Table 6). Nothing is known about the antisense RNA regulation of this transporter in *V.cholerae*. Thus, the antisense RNA regulation might be unique for the *V.anguillarium* system. Although the absence of a Fur-binding site upstream of *fatB* does not allow us to invoke the general rule (Table 1), we still predict the Fur-repressible transcription of the *itsPTUS* operon.

One more iron transport locus in the genome of *Y.pestis* includes a candidate outer membrane siderophore receptor *itrA** and a gene *itrS** (iron transport). We have assigned the function of ItrS* as an iron transport ATPase by sequence

similarity to the ItpS*/vibriobactin protein family. A strong candidate FUR-box was observed upstream of *itrA** (Table 6).

Two *Y.pestis* genes encoding siderophore-biosynthesis proteins, *iucA** and *irp2*, also have candidate FUR-boxes in their upstream regions (Table 6). The former gene is a homolog of the Fur-regulated gene *iucA* from *E.coli* (BLAST *E*-value, 4e-29) and *Shigella flexneri* (BLAST *E*-value, 5e-30). In *E.coli*, the operon *iucABCD* encodes proteins for aerobactin biosynthesis. In *Y.pestis*, *iucA** also belongs to a siderophore-biosynthesis cluster that includes homologs of *iucB*, *iucC* and *iucD* (Fig. 2D). The product of *irp2* participates in yersiniabactin biosynthesis (46) and it lies within a large yersiniabactin biosynthesis and uptake locus with one known and two candidate FUR-boxes (Fig. 2E).

One more siderophore-biosynthesis gene cluster *alcABC** is probably regulated by Fur. The *alcABC** genes lie downstream of genes *fhuF** and *rhsB**, and most likely form an operon with the latter (Fig. 2F). FhuF* is 37% identical to the FhuF protein from *E.coli*. The latter is involved in the reduction of ferric iron in cytoplasmic ferrioxamine B (47). RhsB* is homologous to the siderophore-biosynthesis protein RhsB from *B.subtilis*. A strong FUR-box was found upstream of the *fhuF** gene in *Y.pestis* (Table 6). In the genome of *E.coli*, *fhuF* is also preceded by a strong FUR-box. The genes *alcABC** are homologous to the alcaligin-biosynthesis genes *alcABC* from *B.pertussis* (BLAST *E*-value, e-156). The FUR-box was observed upstream of the *alcABC* operon in the genome of *B.pertussis* and the transcription of *alcABC* is iron dependent (48). Moreover, in the genome of *Y.pestis*, the *fhuF*rhsB*alcABC** operon lies within the locus containing other genes involved in the biosynthesis and transport of siderophores (Fig. 2F). One of them is *alcY**, which is 50% identical to *alcC**. It is also preceded by a candidate FUR-box (Table 6). The downstream gene of *alcY* fauA**, similar to a gene for ferric alcaligin siderophore receptor from *B.pertussis* (BLAST *E*-value, e-147).

A strong candidate FUR-box was found upstream of *psaE*, the transcriptional regulator of biosynthesis of the *Y.pestis* pH6 antigen (Table 6). This antigen plays an essential role in the pathogenesis of *Y.pestis* (49) and the *psa* genes are co-localized with some iron-uptake systems (Fig. 2G). The role of Fur in the regulation of virulence systems is discussed below.

One more candidate Fur-regulated operon with a function in the pathogenesis is *ihaAB** (Table 6). The gene *ihaA** encodes an adhesin similar to the one produced by the pathogenic strain of *E.coli* O157:H7 (*iha*). The latter also has an upstream candidate FUR-box (Table 3; see above). The gene *ihaB** encodes an exogenous ferric siderophore receptor.

The genes encoding coproporphyrinogen oxidase in *Y.pestis* (*chuW**), *E.coli* O157:H7 (*chuW*) and in *V.cholerae* (*VCA0909*) all have candidate FUR-boxes (Table 6). They probably participate in the Fe(III) reduction after it is delivered to the cytoplasm.

Two new *Y.pestis* outer membrane receptors with candidate FUR-boxes are *omrA** and *hasR**. OmrA* is homologous to the members of the ferrisiderophore receptor family including FyuA and FhuA from *E.coli*. The closest homolog of *omrA** was found in the genome of *P.aeruginosa*, where it also has a candidate FUR-box (Table 6). The second gene, *hasR**, is an ortholog of *hasR* from *P.aeruginosa*, encoding a heme receptor. Although the candidate FUR-box of *hasR* in *P.aeruginosa* is weak (Table 6), the Fur-regulated transcription of this gene was shown experimentally (50). In *Y.pestis*, *hasR** lies in a locus containing other iron-acquisition genes (Fig. 2H).

The gene *yebN** encoding a hypothetical membrane protein has a very strong candidate FUR-box. The orthologous gene *yebN* of *E.coli* also has a candidate upstream FUR-box (Table 6). In the *Y.pestis* genome a gene for a ferrichrome receptor follows this gene.

Table 7. The predicted FUR-boxes upstream of the iron-acquisition genes of *V.cholerae* and the FUR-boxes upstream of the homologous genes in other genomes

Gene	Organism	Score	Position	Site
VC1573	VC	5.00	-418	GATAATaATTATCATTtaa
fumC	PA	5.20**	-394	aATAATcAaTcTCATTATC
fumC	EC	No site		
fumC	HI	No site		
*fumC**	ST	4.08	-189	GATAAaGgTAAcgATTtct
*fumC**	YP	4.00	-50	AgaAATGcgAATCAaTcgC
VCA0625	VC	5.00	38	atTAtTGATAATaATTATC
hasR	PA	4.57**	-126	GcaAATGAaAActATTATC
irpA	VC	4.80	-2	acaAATGATAATaATTtgC
irpA	Syn	5.15	-186	aAaAATGATTATtATTcTC
VCA0070	VC	4.67	-269	aAgAATaATTATCgTTATt
VCA0824	VC	4.65	35	tcTtATGcgAATCATTtTC
VCA0232	VC	4.32	-42	atatATGcgAATCgTTATC
frpB	NM	4.14	-131	aAaAAcGATAATCAgctTt
chuA	EC(pat)	5.55	-303	GATAATaATTcTCATTATC
VCA0227	VC	4.15	-2	CtTAATGAgAATaAgTATC
*frpP**	NM	4.10	-83	acTAtaGATTATCATTtat
vibF	VC	5.39	-90	GATAATGATTATtATTAaC)
vibH	VC	4.49	-199	aATAtTGATTcTCATTtcg)
viuP	VC	4.49	10	cgaAATGAgAATCAaTATt)
tonB	VC	4.30	-143	GATAATGcaAtTgATaATt)
vibA	VC	4.25	-94	tcaAATGAgAATagTTtTt)

Genome abbreviation: Syn, Synechococcus; NM, *N.meningitis*; see also Tables 3 and 6.
*Genes that have no common name or have been identified previously as open reading frames only. These genes are named after orthologs from other genomes.
**Experimentally confirmed Fur-binding sites.

Candidate FUR-boxes in the *V.cholerae* genome

Tables 5 and 7 show *V.cholerae* genes with new candidate FUR-boxes: *VC1573*, *VCA0625*, *irpA*, *VCA0070*, *VCA0824*, *VCA0232*, *VCA0227*, *VCA0284*, *VCA0068*, *VC1643*, *VCA0988*, *VCA0923*, *VC1403*, *VC1405*, *tcpI*, *tcpP*, *cheY*, *fliLM*.

VC1573 has a strong candidate FUR-box. It is similar to the fumarate hydratase genes *fumC* from *E.coli* (46% identity), *P.aeruginosa* (65% identity) and other Enterobacteraceae. The transcription of *fumC* from *P.aeruginosa* is regulated by Fur and there is a known Fur-binding site upstream of the *fumC*-containing operon in *P.aeruginosa* (51). However, no FUR-boxes were found upstream of *fumC* in *E.coli* and

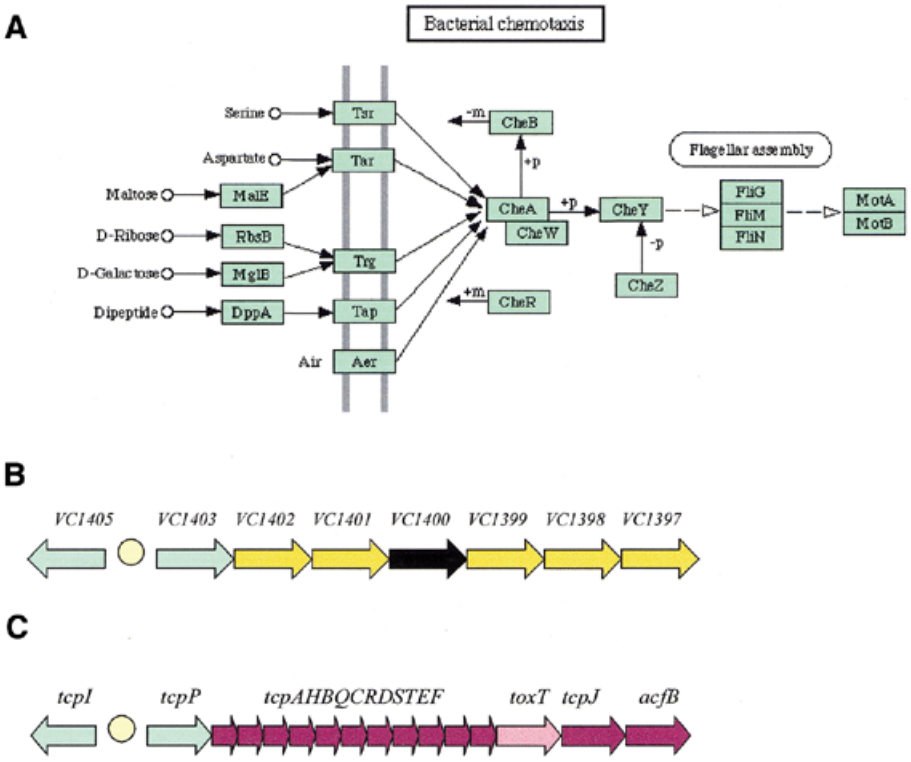

Figure 3. (**A**) The chemotaxis system of *E.coli* (taken from KEGG; 61). (**B**) One of the three loci containing *che* genes. The homologs of *cheW*, *cheB*, *cheR*, *cheY* and *cheA* are shown by the yellow arrows, genes encoding the MCPs are shown by the light green arrows. (**C**) The Tcp-encoding locus. The *tcp* genes similar to MCPs are represented by the light green arrows. All other *tcp* genes are represented by lilac arrows. The gene for the regulator ToxT is shown by the pink arrow.

H.influenzae, and only weak candidate FUR-boxes were observed upstream of *fumC* in *S.typhi* and *Y.pestis* (Table 7).

The product of *VCA0625* was annotated as a TonB receptor-related protein (similar to the heme receptor HasR from *P.aeruginosa* and *Serratia marcescens*). A candidate FUR-box was observed upstream of this gene and upstream of the orthologous gene from *P.aeruginosa* (Table 7).

There is a candidate FUR-box upstream of *VC1264* (Table 7). This gene is similar to *irpA* from *Synechococcus*. In *Synechococcus*, *irpA* is expressed only under iron-deficient conditions and this gene also has a candidate Fur-binding site in the upstream region (Table 7) (52).

VCA0824 encodes a protein similar to diaminobutyrate aminotransferases. This gene has a candidate FUR-box site in the upstream region (Table 7), whereas its product is similar to RhsA, a rhizobactin siderophore-biosynthesis protein from *Sinorhizobium meliloti*.

A new iron-acquisition locus with two candidate FUR-boxes consists of the ABC-transporter operon *VCA0227–VCA0230* and the gene *VCA0232* encoding an outer membrane protein (Fig. 2I). The product of *VCA0232* is similar to FrpB (iron-regulated TonB-dependent outer membrane protein) from *Neisseria meningitis* and ChuA (heme utilization outer membrane protein) from the pathogenic strain of *E.coli* O157:H7. The genes of all three outer membrane receptors have upstream candidate FUR-boxes (Table 7). The Fur-dependent regulation of *chuA* has already been shown (53). The operon *frpPTUS* homologous to the ABC-transport operon *VCA0227–VCA0230* is present in the genome of *N.meningitis* (54) where it lies immediately downstream of *frpB* (Fig. 2J). In both genomes, *V.cholerae* and *N.meningitis*, there are candidate FUR-boxes upstream of the transporter operons (Table 7).

One more iron-acquisition system (*VC0284* and *ybiX**), which is likely to be regulated by Fur, consists of an outer membrane TonB-dependent receptor and an iron-uptake factor (Fig. 2C). This system is similar to the *E.coli* and *P.aeruginosa* iron-acquisition systems (see above).

The profile we used pulled out several *V.cholerae* genes that had already been described as having candidate FUR-box in the upstream regions, namely *vibF* (55), *vibH* (56), *viuP* (56), *tonB* (57) and *vibA* (58).

Unexpectedly, a large set of genes involved in chemotaxis were found to have candidate FUR-boxes in the upstream regions. These genes encode eight MCPs (*VCA0068*, *VC1643*, *VCA0988*, *VCA0923*, *VC1403*, *VC1405*, *tcpI*, *tcpP*); homologs of the *E.coli* chemotactic signal transducing proteins, *VC1397* (*cheA*), *VC1401* (*cheB*), *VC1399* (*cheR*), *VC1402* (*cheW*) and *VC1398* (*cheY*); a flagellar motor switch protein *VC2126* (*fliM*); a homolog of *fliL*, *VC2127*. All mentioned *E.coli* genes are involved in the sequential steps of chemotaxis (Fig. 3A) (26). *Escherichia coli* has only five MCPs: *tar*, *tap*, *air*, *trg*, *tsr* and only one copy of each *che*-gene: *cheA*, *cheW*, *cheR*, *cheB*, *cheY*, *cheZ*. On the contrary, more than 40 MCPs were found in the genome of *V.cholerae* and there are three large loci with *che*-like genes. A candidate FUR-box was observed upstream of the locus containing genes least similar to the *E.coli* *che*-genes (Fig. 3B). Table 8 shows candidate FUR-boxes

Table 8. The candidate FUR-boxes upstream of genes encoding MCPs in the *V.cholerae* genome

Gene	Score	Position	Site
VCA0068	4.67	-44	AATAAcGATAATtATTcTt
VC1643	4.44	-132	GATAgTGATgATaATTAaC
VCA0988	4.34	-12	GAcAAcGATAATCAcaATC
VCA0923	4.39	-6	tATAAatcaAATCATTATC
VC1403	4.32	-105	GAaAATttaAATCATTgTC
VC1405	4.32	-354	GAcAATGATTtaaATTtTC
tcpI	4.52	-272	GAaAgTaAATcTCATTATa
tcpP	4.52	-165	tATAATGAgAATtAcTtTC
tcpJ	4.48	-134	GAatATcATAtTCAaTtTC
fliLM	4.26	46	acTgtTGATcATCATTATC

Table 9. Sequences similar to FUR-boxes upstream of genes with no direct function in iron acquisition

Gene	Organizm	Score	Position	Site
gpmA	EC	5.28	-46	tATAATGAgAATtATTATC
*gpmA**	ST	5.28	-24	TATAATGAgAATtATTATC
*gpmA**	YP	5.73	-185	tATAATGATAATaATTATC
nupC	EC	5.06	-287	GAgAATGATTATCAaatTC
*nupC**	ST	4.82	-294	GATAATGATTATCAagtgC
*nupC**	YP	4.97	-382	ttaAATGAgAATaATTATC
b2392	EC	5.06	-67	GAattTGATAATCATTcTC
*NRAMP**	ST	4.82	-48	GcactTGATAATCATTATC
*NRAMP**	YP	4.97	-70	GATAATtATTcTCATTtaa

*Genes that have no common name or have been identified previously as open reading frames only. These genes are named after orthologs from other genomes.

upstream of *V.cholerae* MCPs, *che*-like genes and *fliLM* homologs. The chemotaxis plays an important part in the *V.cholerae* pathogenesis (59). Indeed, two of the MCPs with candidate FUR-boxes (*tcpI* and *tcpP*) are the members of the toxin co-regulated pilus-biosynthesis cluster (Fig. 3C). The role of Fur in the pathogenesis is discussed below (see Discussion).

Some exceptions

The profiles and comparison-based filters produced three more genes with conserved Fur-box-like sequences in the upstream regions, but the functions of these genes were not linked to the iron metabolism (Table 9). Thus, they seem to be false positives. These genes are: *gpm*A (phosphoglycerate mutase 1 involved in glycolysis), nupC (nucleoside permease) and *b2392* (NRAMP manganese transport protein) in the genome of *E.coli* (both strains) and their orthologs in the genomes of *S.typhi* and *Y.pestis*.

DISCUSSION

This study identified a number of likely Fur-regulated genes in the genomes of *E.coli*, *S.typhi*, *Y.pestis* and *V.cholerae*. Most of these genes either are known to function in iron acquisition and metabolism or exhibit significant homology to proteins of such systems. The presence of FUR-boxes upstream of the latter genes allows us to ascribe the function in iron metabolism to them with higher confidence. Thus, a number of new iron-acquisition mechanisms were described, although this annotation still requires an experimental confirmation. Here we describe the role of Fur in the biosynthesis of the new siderophores and the regulation of the ferrisiderophore-uptake pathways; the fate of the Fe(III) ion taken into the cytoplasm, the regulation of the enzymes involved in the Fe(III) reduction and protection of the free iron. Finally, we discuss the role of Fur as a global regulator of pathogenesis.

Siderophore biosynthesis

Several new siderophore biosynthesis genes were found in the genomes of *Y.pestis* and *V.cholerae*. In the genome of *Y.pestis* there are two new loci. One of them contains genes similar to the aerobactin-biosynthetic operon of *E.coli* and *S.flexneri*, whereas the other contains genes similar to the alcaligin-biosynthetic operon of *B.pertussis* (Fig. 2D and F). Both gene clusters are preceded by candidate FUR-boxes and their transcription is likely to be Fur regulated. *Vibrio cholerae* has a gene homologous to *rhsA* from *S.meliloti* that encodes the first enzyme of the rhizobactin-biosynthesis pathway. It has also an upstream candidate FUR-box and, given its function, is likely to be under the regulation of Fur.

Iron-uptake systems

A new family of iron-acquisition systems was identified in *E.coli* (*fptAybiX*) and *V.cholerae* (*VCA0284*, *ybiX**). Each system consists of a TonB-dependent outer membrane receptor (*fptA*, *VCA0284*) and an iron-uptake factor (*ybiX*, *ybiX**). A similar system was found in *P.aeruginosa* (*piuAC*), where it is known to be regulated by Fur (33) (Fig. 2A–C). So we propose Fur regulation for the *E.coli* and *V.cholerae* systems as they also have candidate FUR-boxes.

Three new Fe(III)-ABC transporters of *Y.pestis* (*itpPTS**, *itsPTUS** and *itrAS**) were identified by sequence similarity to iron-transport systems of other bacteria. Their regulation by Fur was predicted based on the following observation: the presence of FUR-boxes, the localization of the genes within large iron-acquisition loci and, in the case of *its**, the conservation of the site in the genome of *V.cholerae*.

A new hemin-uptake system probably regulated by Fur is described in *E.coli* and *S.typhi*. In the genome of *E.coli*, it is encoded by two separately located genes *b1995* and *ydiE*. Each of them is preceded by a candidate FUR-box. In the genome of *Y.enterocolitica* these genes, *hemP** and *hemR**, are immediately adjacent and transcribed in the same direction. A similar system has already been identified in *Y.pestis*, where these two genes are divergently transcribed from Fur-dependent promoters (34).

One likely Fur-regulated Fe(III) transport system is described in the genome of *V.cholerae*. It includes the outer membrane receptor *VCA0232* and the ABC transporter *VCA0227–VCA0231*. Both *VCA0232* and *VCA0227* are preceded by candidate FUR-boxes. A similar system was found in *N.meningitis* (*frpB* and *frpPTUS**). Both *frpB* and *frpP** are preceded by candidate FUR-boxes.

Three new outer membrane ferrisiderophore and hemin receptors seem to be under Fur regulation in the genomes of *Y.pestis* and *V.cholerae*: *omrA**, *hasR** (*Y.pestis*) and *VCA0625* (*V.cholerae*).

Fe(III) release

As Fe(II) is bound by siderophores weaker than Fe(III), the enzymatic reduction of iron as a release mechanism has been proposed (1). We predict that this function is encoded by the oxidoreductase gene *yhhX* from *E.coli*, because this gene has candidate FUR-boxes in the genomes of *E.coli*, *S.typhi* and *B.subtilis*.

Coproporphyrine oxidases from *Y.pestis* (*hemN**) and *V.cholerae* (*VCA0909*) also have candidate FUR-boxes in both genomes and thus they are likely to be regulated by Fur.

The gene *fhuF** has FUR-boxes in *Y.pestis* and *E.coli* and it is probably involved in the reduction of the ferric iron in the cytoplasmic ferrisiderophores.

We also propose Fur regulation for the *E.coli* gene *b3070*, whose product is homologous to the iron-chelator utilization protein ViuB of *V.cholerae*.

Superoxide dismutase

Manganese-containing superoxide dismutase, encoded by the gene *sodA*, plays a crucial role in maintaining low intracellular levels of the superoxide radical (O_2^-). The superoxide can directly cause cell damage. Besides, the reduction of iron by the superoxide and the presence of H_2O_2 provides the reactants for the Fenton reaction leading to the formation of the hydroxyl radicals that can attack all biological macromolecules. Thus, the expression of *sodA* should be strictly controlled. In *E.coli* it is repressed by Fur, ArcA, Fnr and integration host factor IHF, and it is activated by the product of the *soxRS* locus and mutated alleles of *soxQ* locus (such as CfxB). As FUR-boxes are found upstream of the orthologous genes in *S.typhi*, *Y.pestis* and *V.cholerae*, we predict that the Fur regulation of this gene is conserved.

The role of Fur in pathogenesis

The concept of Fur as a global regulator of pathogenesis processes has already been proposed. Indeed, as in the host environment iron is rarely present in the free form, the low level of exogenous iron can indicate that the invasion has occurred. Thus, it is natural to expect some of pathogenesis mechanisms to be controlled directly by Fur. For example, in Stoebner and Payne (61) it was shown that *V.cholerae* produces a Fur-regulated hemolysin, which is involved in the acquisition of iron from heme or hemoglobin. In this study we have observed four possible examples of Fur regulation of virulence systems.

The gene *iha* encoding adhesin in *Y.pestis* and *E.coli* O157:H7 is likely to be Fur regulated as it has upstream FUR-boxes in both genomes.

The gene *Z2267* in *E.coli* O157:H7 encoding a putative receptor with a candidate role in the enterotoxin production also has a candidate FUR-box. It is conserved in the genome of *S.typhi*.

The *psaE* gene of *Y.pestis* encodes a regulator of the *Y.pestis* pH6 antigen biosynthesis. Again, a candidate FUR-box was observed upstream of *psaE*.

Chemotaxis plays an essential role in the pathogenesis of *V.cholerae*. We have detected FUR-boxes upstream of genes encoding proteins involved in all sequential steps of the chemotaxis signal transduction: (i) eight MCP-encoding genes, (ii) an operon homologous to the signal transducing *che*-system and (iii) the gene *fliM* encoding a flagellar motor switch protein. Thus, we propose that the pathogenic chemotaxis of *V.cholerae* is induced in response to the iron-restricted environments facing the bacterium within the human body, and Fur is the key regulator that mediates this response. Moreover, the TCP gene cluster of *V.cholerae* encoding the toxin-coregulated pilus also has a FUR-box and thus can be induced at the time of invasion in the host.

CONCLUSION

One usually thinks of horizontal gene transfer as a simultaneous transfer of an operon and a transcription regulatory factor. The perfect example is the transfer of *trpR* and *trpAB* in *Chlamidia trachomatis* (15) or numerous plasmid-encoded loci. The reason is that in most cases, original regulators of homologous systems diverged significantly in distant species. For example, the orthologous factors LexA/DinP or the phosphate repressors have absolutely different binding sites in Gram-positive and Gram-negative bacteria (62). The arginine repressor ArgR/AhrC has similar recognition signals in Gram-positive and Gram-negative species, but the cross talk is almost impossible (63). The situation with the Fur repressor is different. As iron acquisition is closely connected to the processes of pathogenesis, Fur-dependent systems are a frequent subject of horizontal gene transfer within pathogenic islands, virulence plasmids, etc. That is why Fur protein itself, as well as its recognition signal, is highly conserved among diverse species. As a result, in the case of Fur we frequently observe transfer of systems without regulators from one bacterium to another. Then these systems fall under regulation by pre-existing factors that bind to the same FUR-boxes as in the source sequences. However, we should note that there is at least one example of a system where the signal is highly conserved with no evidence of horizontal transfer; it is the CIRCE element bound by the heat shock regulator HrcA (65).

One more consequence of the frequent horizontal transfer of iron-acquisition systems is that they usually form large loci, sometimes with several Fur-binding sites within them. It is interesting to follow the fate of such loci in various species: however close they are, the arrangement of genes is usually totally different. Some genes disappear, and instead new ones are included. Sometimes even the place of Fur-binding site changes.

Most bacteria have only one representative of the Fur family encoded in the genome, but there are still some exceptions: in the genome of *B.subtilis* there are three paralogs: *fur* (ferrum-uptake regulator), *zur* (zinc-uptake regulator) and *perR* (peroxide-regulon repressor). Besides, some Fur-regulated genes are also under the control of other regulators. For example, the iron-regulated protein A (*irgA*) in *V.cholerae* is repressed by Fur and activated by IrgB, the yersiniabactin biosynthesis genes are under negative control by Fur and YbtA. In this work we have found a similar system with an 'additional' regulator in *S.typhi*. One more interesting example of additional regulation is present in *E.coli* K12. It is the FecIR

system. FecI is a specific sigma factor that is used for the transcription of the *fecABCDE* genes (ferridicitrate-transport system). This system was totally lost in the genomes of *E.coli* O157:H7, *S.typhi*, *Y.pestis* and *V.cholerae*, but the regulatory system FecI/FecR and the receptor gene *fecA* are present in several copies in the genome of *P.aeruginosa* (E.M.Panina and M.S.Gelfand, study in progress).

The role of Fur in pathogenesis is of particular interest. There are many papers reporting experiments where mutations in some Fur-dependent iron-transport genes prevented the successive invasion by the pathogen. Here we show not only that the iron-acquisition systems are Fur-controlled, but also that some of the direct virulence determinants such as adhesins or toxin-production genes probably are under Fur control. Although the role of Fur in pathogenesis is a subject of an independent study now in progress (E.M.Panina and M.S.Gelfand), even at this stage it is clear that Fur can be a promising target for drug design.

ACKNOWLEDGEMENTS

We are grateful to Alexandra B. Rakhmaninova, Olga Laykova and Dmitry Rodionov for useful discussions. This study was partially supported by grants from INTAS (99-1476) and HHMI (55000309).

REFERENCES

1. Grunberg-Manago,M. (1996) Regulation of the expression of aminoacyl-tRNA synthetases and translation factors. In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella. Cellular and Molecular Biology*. ASM Press, Washington DC, pp. 1075–1090.
2. Braun,V. and Hantke,K. (1991) Genetics of bacterial iron transport. In Winkelmann,G. (ed.), *CRC Handbook of Microbial Iron Chelate*. CRC Press Inc., Boca Raton, FL, pp. 107–130.
3. Fecker,L. and Braun,V. (1983) Cloning and expression of the *fhu* genes involved in iron(III)-hydroxamate uptake by *Escherichia coli*. *J. Bacteriol*., **156c**, 1301–1314.
4. van Hove,B., Staudenmaier,H. and Braun,V. (1990) Novel two-component transmembrane transcription control: regulation of iron dicitrate transport in *Escherichia coli* K-12. *J. Bacteriol*., **172**, 6749–6758.
5. Sauer,M., Hantke,K. and Braun,V. (1990) Sequence of the *fhuE* outer membrane receptor gene of *Escherichia coli* K-12 and properties of mutants. *Mol. Microbiol.*, **4**, 427–437.
6. Earhart,C. (1987) Ferrienterobactin transport in *Escherichia coli.* In Winkelmann,G., Van der Helm,D. and Neilands,J.,B. (eds), *Iron Transport in Microbes*, *Plants and Animals*. VCH Verlagsgesellschaft, Weinheim, Germany, pp. 67–84.
7. Klebba,P.E., Rutz,J.M., Liu,L. and Murphy,C.K. (1993) Mechanisms of TonB-catalyzed iron transport through the entire bacterial cell envelope. *J. Bioenerg. Biomembr.*, **25**, 603–611.
8. Halliwell,B. and Gutteridge,M.C. (1984) Oxygen toxicity, oxygen radicals, transition metals and decease. *Biochem. J.*, **219**, 1–14.
9. Stojiljkovic,I., Baumler,A.J. and Hantke,K. (1994) Fur regulon in Gram-negative bacteria. Identification and characterization of new iron-regulated *Escherichia coli* genes by a Fur titration assay. *J. Mol. Biol.* **236**, 531–545.
10. de Lorenzo,V., Herrero,M., Giovannini,F. and Neilands,J.B. (1988) Metal ion regulation of gene expression. Fur repressor-operator interaction at the promoter region of the aerobactin system of pColV-K30. *J. Mol. Biol.*, **203**, 875–884.
11. Griggs,D.W. and Konisky,J. (1989) Mechanism for iron-regulated transcription of the *Escherichia coli* cir gene: metal-dependent binding of Fur protein to the promoters. *J. Bacteriol.*, **171**, 1048–1054.
12. Mironov,A.A., Koonin,E.V., Roytberg,M.A. and Gelfand,M.S. (1999) Computer analysis of transcription regulatory patterns in completely sequenced bacterial genomes. *Nucleic Acids Res.*, **27**, 2981–2989.
13. Gelfand,M.S., Novichkov,P.S., Novichkova,E.S. and Mironov,A.A. (2000) Comparative analysis of regulatory patterns in bacterial genomes. *Brief. Bioinform.*, **1**, 357–371.
14. Gelfand,M.S., Koonin,E.V. and Mironov,A,A. (2000) Prediction of transcription regulatory sites in Archaea by a comparative genomic approach. *Nucleic Acids Res.*, **28**, 695–705.
15. Panina,E.M., Vitreschak,AG., Mironov,A.A. and Gelfand,M.S. (2001) Regulation of aromatic amino acid biosynthesis in pamma-proteobacteria. *J. Mol. Microbiol. Biotechnol.*, **3**, 529–543.
16. Rodionov,D.A., Mironov,A.A., Rakhmaninova,A.B. and Gelfand,M.S. (2000) Transcriptional regulation of transport and utilization systems for hexuronides, hexuronates and hexonates in gamma purple bacteria. *Mol. Microbiol.*, **38**, 673–683.
17. McGuire,A.M., Hughes,J.D. and Church,G.M. (2000) Conservation of DNA regulatory motifs and discovery of new motifs in microbial genomes. *Genome Res.*, **10**, 744–757.
18. McCue,L.A., Thompson,W., Carmack,C.S., Ryan,M.P., Liu,J.S., Derbyshire,V. and Lawrence,C.E. (2001) Phylogenetic footprinting of transcription factor binding sites in proteobacterial genomes. *Nucleic Acids Res.*, **29**, 774–782.
19. Zhou,D., Hardt,W.D. and Galan,J.E. (1999) *Salmonella typhi* encodes a putative iron transport system within the centisome 63 pathogenicity island. *Infect. Immun.*, **67**, 1974–1981.
20. Chai,S., Welch,T.J. and Crosa,J.H. (1998) Characterization of the interaction between Fur and the iron transport promoter of the virulence plasmid in *Vibrio anguillarum*. *J. Biol. Chem.*, **273**, 33841–33847.
21. Bsat,N., Herbig,A., Casillas-Martinez,L., Setlow,P. and Helmann,J.D. (1998) *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. *Mol. Microbiol.*, **29**, 189–198.
22. Thomas,C.E. and Sparling,P.F. (1996) Isolation and analysis of a fur mutant of *Neisseria gonorrhoeae*. *J. Bacteriol.*, **178**, 4224–4232.
23. Xiong,A., Singh,V.K., Cabrera,G. and Jayaswal,R.K. (2000) Molecular characterization of the ferric-uptake regulator, Fur, from *Staphylococcus aureus*. *Microbiology*, **146**, 659–668.
24. Rowland,B.M., Grossman,T.H., Osburne,M.S. and Taber,H.W. (1996) Sequence and genetic organization of a *Bacillus subtilis* operon encoding 2,3-dihydroxybenzoate biosynthetic enzymes. *J. Bacteriol.*, **178**, 119–123.
25. Litwin,C.M. and Calderwood,S.B. (1993) Role of iron in regulation of virulence genes. *Clin. Microbiol. Rev.* **6**, 137–149.
26. Grunberg-Manago,M. (1996) Chemotaxis. In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella. Cellular and Molecular Biology*. ASM Press, Washington DC, pp. 1103–1126.
27. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J., Rapp,B.A. and Wheeler,D.L. (2000) GenBank. *Nucleic Acids Res.*, **28**, 15–18.
28. Schneider,T.D., Stormo,G.D., Gold,L. and Ehrenfeucht,A. (1986) Information content of binding sites on nucleotide sequences. *J. Mol. Biol.*, **188**, 415–431.
29. Peredo,A., Saint-Pierre,C., Latour,J.M., Michaud-Soret,I. and Forest,E. (2001) Conformational changes of the ferric uptake regulation protein upon metal activation and DNA binding; first evidence of structural homologies with the diphtheria toxin repressor. *J. Mol. Biol.*, **310**, 83–91.
30. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876–4882.
31. Mironov,A.A., Vinokurova,N.P. and Gelfand,M.S. (2000) GenomeExplorer: software for analysis of complete bacterial genomes. *Mol. Biol.*, **34**, 222–231.
32. Oshima,T., Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Ikemoto,K., Inada,T., Itoh,T., Kajihara,M. *et al.* (1996) A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7–28.0 min region on the linkage map. *DNA Res.*, **3**, 137–155.
33. Ochsner,U.A. and Vasil,M.L (1996) Gene repression by the ferric uptake regulator in *Pseudomonas aeruginosa*: cycle selection of iron-regulated genes *Proc. Natl Acad. Sci. USA*, **93**, 4409–4414.
34. Stojiljkovic,I. and Hantke,K. (1992) Hemin uptake system of *Yersinia enterocolitica*: similarities with other TonB-dependent systems in gram-negative bacteria. *EMBO J.*, **11**, 4359–4367.
35. Butterton,J.R., Choi,M.H., Watnick,P.I., Carroll,P.A. and Calderwood,S.B. (2000) *Vibrio cholerae* VibF is required for vibriobactin synthesis and is a member of the family of nonribosomal peptide synthetases. *J. Bacteriol.*, **182**, 1731–1738.

36. Butterton,J.R. and Calderwood,S.B. (1994) Identification, cloning and sequencing of a gene required for ferric vibriobactin utilization by *Vibrio cholerae*. *J. Bacteriol.*, **176**, 5631–5638.
37. Vassimiva,N. and Kozyrev,D. (2000) A method for direct cloning of Fur-regulated genes: identification of seven new Fur-regulated loci in *Escherichia coli*. *Microbiology*, **146**, 3171–3182.
38. Tarr,P.I., Bilge,S.S., Vary,J.C.,Jr, Jelacic,S., Habeeb,R.L., Ward,T.R., Baylor,M.R. and Besser,T.E. (2000) Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infect. Immun.*, **68**, 1400–1407.
39. Jones,H.A., Lillard,J.W. and Perry,R.D. (1999) HmsT, a protein essential for expression of the haemin storage (Hms+) phenotype of *Yersinia pestis*. *Microbiology*, **145**, 2117–2128.
40. Pradel,E., Guiso,N. and Locht,C. (1998) Identification of AlcR, an AraC-type regulator of alcaligin siderophore synthesis in *Bordetella bronchiseptica* and *Bordetella pertussis*. *J. Bacteriol.*, **180**, 871–880.
41. Fetherston,J.D., Bearden,S.W. and Perry,R.D. (1996) YbtA, an AraC-type regulator of the *Yersinia pestis* pesticin/yersiniabactin receptor. *Mol. Microbiol.*, **22**, 315–325.
42. Heinrichs,D.E. and Poole,K. (1993) Cloning and sequence analysis of a gene (pchR) encoding an AraC family activator of pyochelin and ferripyochelin receptor synthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.*, **175**, 5882–5889.
43. Tsolis,R.M., Baumler,A.J., Stojiljkovic,I. and Heffron,F. (1995) Fur regulon of *Salmonella typhimurium*: identification of new iron-regulated genes. *J. Bacteriol*., **177**, 4628–4637.
44. Actis,L.A., Tolmasky,M.E., Crosa,L.M. and Crosa,J.H. (1995) Characterization and regulation of the expression of FatB, an iron transport protein encoded by the pJM1 virulence plasmid. *Mol. Microbiol.* **17**, 197–204.
45. Chen,Q. and Crosa,J.H. (1996) Antisense RNA, fur, iron and the regulation of iron transport genes in *Vibrio anguillarum*. *J. Biol. Chem.*, **271**, 18885–18889.
46. Gehring,A.M., DeMoll,E., Fetherston,J.D., Mori,I., Mayhew,G.F., Blattner,F.R., Walsh,C.T. and Perry,R.D. (1998) Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*. *Chem. Biol.*, **50**, 573–586.
47. Muller,K., Matzanke,B.F., Schunemann,V., Trautwein,A.X. and Hantke,K. (1998) FhuF, an iron-regulated protein of *Escherichia coli* with a new type of [2Fe-2S] center. *Eur. J. Biochem.*, **258**, 1001–1008.
48. Giardina,P.C., Foster,L.A., Toth,S.I., Roe,B.A. and Dyer,D.W. (1997) Analysis of the alcABC operon encoding alcaligin biosynthesis enzymes in *Bordetella bronchiseptica*. *Gene*, **194**, 19–24.
49. Lindler,L.E., Klempner,M.S. and Straley,S.C. (1990) *Yersinia pestis* pH 6 antigen: genetic, biochemical and virulence characterization of a protein involved in the pathogenesis of bubonic plague. *Infect. Immun.*, **58**, 2569–2577.
50. Ochsner,U.A., Johnson,Z. and Vasil,M.L. (2000) Genetics and regulation of two distinct haem-uptake systems, phu and has, in *Pseudomonas aeruginosa*. *Microbiology*, **146**, 185–198.
51. Ochsner,U.A., Johnson,Z. and Vasil,M.L. (1997) An operon containing fumC and sodA encoding fumarase C and manganese superoxide dismutase is controlled by the ferric uptake regulator in *Pseudomonas aeruginosa*: fur mutants produce elevated alginate levels. *J. Bacteriol*., **179**, 1452–1459.
52. Reddy,K.J., Bullerjahn,G.S., Sherman,D.M. and Sherman,L.A. (1988) Cloning, nucleotide sequence and mutagenesis of a gene (irpA) involved in iron-deficient growth of the cyanobacterium Synechococcus sp. strain PCC7942. *J. Bacteriol.*, **170**, 4466–4476.
53. Torres,A.G. and Payne,S.M. (1997) Haem iron-transport system in enterohaemorrhagic *Escherichia coli* O157:H7. *Mol. Microbiol.*, **23**, 825–833.
54. Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J. *et al*. (2000) Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. *Science*, **287**, 1809–1815.
55. Butterton,J.R., Choi,M.H., Watnick,P.I., Carroll,P.A. and Calderwood,S.B. (2000) *Vibrio cholerae* VibF is required for vibriobactin synthesis and is a member of the family of nonribosomal peptide synthetases. *J. Bacteriol.*, **182**, 1731–1738.
56. Wyckoff,E.E., Valle,A.M., Stacey,L.S. and Payne,S.M (1999) A multifunctional ATP-binding cassette transporter system from *Vibrio cholerae* transports vibriobactin and enterobactin. *J. Bacteriol*., **181**, 7588–7596.
57. Occhino,D.A., Wyckoff,E.E., Henderson,D.P., Wrona,T.J. and Payne,S.M. (1998) *Vibrio cholerae* iron transport: haem transport genes are linked to one of two sets of tonB, exbB, exbD genes. *Mol. Microbiol.*, **29**, 1493–1507.
58. Wyckoff,E.E., Stoebner,J.A., Reed,K.E. and Payne,S.M. (1997) Cloning of a *Vibrio cholerae* vibriobactin gene cluster: identification of genes required for early steps in siderophore biosynthesis. *J. Bacteriol*., **179**, 7055–7062.
59. Freter,R., O'Brien P.C. and Macsai,M.S. (1979) Effect of chemotaxis on the interaction of cholera vibrios with intestinal mucosa. *Am. J. Clin. Nutr.*, **32**, 128–132.
60. Stoebner,J.A. and Payne,S.M. (1998) Iron-regulated hemolysin production and utilization of heme and haemoglobin by *Vibrio cholerae*. *Infect. Immun.*, **56**, 2891–2895.
61. Ogata,H., Goto,S., Sato,K., Fujibuchi,W., Bono,H. and Kanehisa,M. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **27**, 29–34.
62. Permina,E.A., Mironov,A.A. and Gelfand,M.S. (2000) Comparative approach to analysis of regulation in complete genomes: SOS repair in *Bucillus subtilis*. Proceeding of the 2nd International Conference on Bioinformatics of Genome Regulation and Structure, **2**, 22–23.
63. Makarova,K.S., Mironov,A.A. and Gelfand,M.S. (2001) Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biol.*, **2**, research0013.1–0013.8.
64. Gelfand,M.S. (1999) Recognition of regulatory sites by genomic comparison. *Res. Microbiol.*, **150**, 755–771.