*Microbial Pathogenesis* 1998; **25**: 77–90
Article No. mi980217

MICROBIAL
PATHOGENESIS

# Identification of PhoP-PhoQ activated genes within a duplicated region of the *Salmonella typhimurium* chromosome

**John S. Gunn[a]*, William J. Belden[b] & Samuel I. Miller[a]†**

*University of Washington, Departments of Medicine and Microbiology, HSB K-140, Box 357710, Seattle, WA 98195, U.S.A., Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire, 03755, U.S.A.*

(Received February 16, 1998; accepted in revised form May 14, 1998)

Salmonellae virulence requires the PhoP-PhoQ two-component regulatory system. PhoP-PhoQ activate the transcription of genes following phagocytosis by macrophages which are necessary for survival within the phagosome environment. Thirteen previously undefined PhoP-activated gene fusions generated by Mu*dJ* and Tn*phoA* (*pag A*, and *E-P*, respectively) were cloned and sequenced. Most *pag* products show no similarity to proteins in the database, while others are predicted to encode: a UDP-glucose dehydrogenase (*pagA*); a protein with similarity to the product of an *E. coli* aluminium-induced gene (*pagH*); a protein encoded within a *Salmonella*-unique region adjacent to the *sinR* gene (*pagN*); a protein similar to a product of the *Yersinia* virulence plasmid (*pagO*); and a protein with similarity to CrcA which is necessary for resistance of *E. coli* to camphor (*pagP*). Of the *pag* characterized, only *pagK, M* and *O* were closely linked. *pagJ* and *pagK* were shown to be unlinked but nearly identical in DNA sequence, as each was located within a 1.6 kb DNA duplication. The translations of sequences surrounding *pagJ* and *pagK* show similarity to proteins from extrachromosomal elements as well as those involved in DNA transposition and rearrangement, suggesting that this region may have been or is a mobile element. The transcriptional start sites of *pagK, M*, and *J* were determined; however, comparison to other known *pag* gene promoters failed to reveal a consensus sequence for PhoP-regulated activation. DNA sequences hybridizing to a *Salmonella typhimurium pagK* specific probe were found in *S. enteritidis* but absent in other *Salmonella* serotypes and Enterobacteriaceae tested, suggesting that these genes are specific for broad host range Salmonellae that cause diarrhoea in humans. Cumulatively, these data further demonstrate: (1) that PhoP-PhoQ is a global regulator of the production of diverse envelope or secreted proteins; (2) that PhoP-PhoQ regulate the production of proteins of redundant function; and (3) that *pag* are often located in regions of horizontally acquired DNA that are absent in other Enterobacteriaceae. 

**Key words:** *Salmonella typhimurium*, PhoP-PhoQ, *pag*, virulence, DNA duplication.

* Current address: University of Texas Health Science, Center at San Antonio, Department of Microbiology, 7703 Floyd Curl Drive, San Antonio, TX 78284-7758, U.S.A.
† Author for correspondence.

0882–4010/98/080077 + 14 $30.00/0

## Introduction

Salmonellae are enteric pathogens that cause a wide range of host-specific illnesses, including gastroenteritis and enteric fever. *Salmonella typhimurium* and *S. typhi* virulence is, in part, controlled by the two-component regulatory system, PhoP-PhoQ [1, 2]. PhoP-PhoQ mutants have markedly reduced virulence for humans (*S. typhi*) and mice (*S. typhimurium*) [1, 3, 4], have decreased ability to survive within cultured macrophages [1, 3], are more sensitive to the action of numerous antimicrobial peptides [4-7], are more efficiently processed and presented by macrophages (as assessed by strain carrying a heterologous antigen) [8] and are unable to induce lipid A modifications in response to environmental signals [9].

PhoQ is a membrane bound sensor-kinase [10] that specifically phosphorylates PhoP, and PhoP-phosphate activates or represses the transcription of target genes. PhoP-PhoQ are thought to positively and negatively regulate the production of over 40 gene products [11]. Genes activated by PhoP-PhoQ (*pag*) have increased expression after bacteria are phagocytosed by macrophages [12]. The specific host factors that activate PhoQ are unknown, however low levels of $Mg^{2+}$ within the phagosome have been suggested to be these factors, as the divalent cations $Ca^{2+}$ and $Mg^{2+}$ repress *pag* expression and bind the periplasmic domain of PhoQ [13-15].

Several PhoP-activated and PhoP-repressed (*prg*) loci have been identified. Database and functional similarities to products of PhoP-repressed genes include proteins that are components or targets of the SPI1 Type III secretion system [16, 17]. PhoP-activated genes include *mgtA* and *mgtCB*, encoding two high-affinity magnesium transport systems [15, 18]; *phoP*, which is autoregulated [19]; *pmrAB*, a two-component system governing resistance to antimicrobial peptides including polymyxin [20, 21]; *phoN*, a periplasmic non-specific acid phosphatase [22]; and outer membrane and envelope proteins including those produced by *pagC* and *pagD* [23, 24]. Other *pag* have been identified that exhibit similarity to an open reading frame (orf) located near the *E. coli mar* locus; an orf with similarity to VanX, a dipeptidase required for resistance to vancomycin; and an orf upstream of the *nhoA* gene of *S. typhimurium* [25]. Through the use of IVET, other *pag* have also been identified [26]. Although strains with a *phoP* or *phoQ* null mutation have numerous observable phenotypic alterations such as attenuation of virulence and reduced antimicrobial peptide resistance, no individual PhoP-activated gene deletion has been shown to result in the identical magnitude of any of the phenotypes of a PhoP-null mutant. Therefore, it is possible that *pag* encode proteins of redundant function. In this work, we reveal information derived from sequence analysis of a number of *pag*, and present data demonstrating the genetic redundancy of PhoP-activated genes.

## Results

### Characterization of PhoP-activated loci

We have previously reported the cloning and characterization of *pagB* [20], *pagC* [1, 24], and *pagD* [23, 27], and the identification of *pagA* and *pagE* to *pagP* by transposon mutagenesis [1, 23]. To understand better the PhoP-PhoQ regulon and the function of genes activated by this regulon, the DNA adjacent to the transposon insertions defining *pagA* and *pagE* to *pagP* was cloned and sequenced with primers homologous to the ends of Mu*dJ* or Tn*phoA*. This sequence was then used to search the database for similar genes or proteins. Database similarities were identified for *pagA, pagH, pagN*, *pagO*, and *pagP*. The predicted PagA protein is similar to UDP-glucose dehydrogenases found in several organisms including *Shigella flexneri, E. coli, Vibrio cholerae* and *Streptococcus pneumoniae* [28–31]. Transcription of *pagA* is activated in response to acidic conditions of the growth medium [12, 21, 32] and is directly regulated by the PmrA-PmrB two-component system [21, 33]. PagA (now referred to as PmrE) is necessary for both polymyxin resistance and for the addition of aminoarabinose to lipid A [33]. It was recently shown that PhoP-PhoQ activated transcription of *pmrA-pmrB* [20, 21], which explains why this gene was originally identified as a *pag*.

*pagN* is located adjacent to *sinR*, encoding a putative transcriptional activator that is located in a chromosomal region unique to *Salmonella* spp. but absent in other enterics [34]. *pagN* is 3′ to *sinR* and transcribed in the opposite direction. A small orf exists between *pagN* and *sinR*, and a second orf exists upstream of *pagN*. These orfs are transcribed in the same direction as *pagN*, but it is unknown if any of these putative genes

are co-transcribed. The predicted *pagN* product is similar in protein sequence to that of the *tia* invasion gene of enteropathogenic *E. coli* [35]. The *tia* gene encodes a protein that, when expressed from a multicopy plasmid, allows *E. coli* DH5α to enter epithelial cells.

The predicted translated product of *pagH* shows similarity to the Ais protein of *E. coli*, the production of which is induced by aluminium [36]. Like *pagA (pmrE), pagH* has been shown to be directly regulated by PmrA-PmrB and not PhoP-PhoQ [33]. *pagO* encodes a protein similar to an open reading frame located downstream of *yadA* encoded on the *Yersinia* virulence plasmid [37], as well as uncharacterized orfs of *Buchnera aphidicola* and *Clostridium kluyveri*. Finally, *pagP* encodes a protein with similarity to the predicted protein of the *E. coli* gene *crcA*, involved in resistance to camphor [38]. Details of these similarities are described in Table 1.

## *pagK*, *M* and *O* are linked on the chromosome

Comparison between the *pag* DNA sequences showed identity within the regions extending from the transposon insertions defining *pagK, pagM* and *pagO*. Southern blot experiments using probes corresponding to *pagK, pagM* and *pagO* showed identical hybridization patterns with *S. typhimurium* chromosomal DNA digested with *Eco*RV or *Hin*dIII, suggesting that these loci are linked on the chromosome (data not shown). The DNA sequence of this region was derived, and all three genes were found to reside within a span of 3.8 kb (Fig. 1). *pagK* should encode a protein of 66 amino acids (7.6 kDa, predicted) and, consistent with its identification by Tn*phoA* mutagenesis, contains a typical, hydrophobic bacterial signal sequence at the amino terminus [39]. *pagM* is predicted to encode a small molecular weight protein (60 amino acids; 6.1 kDa, predicted) and contains a signal sequence at its amino terminus. Each gene is transcribed from its own promoter since the genes are divergently transcribed and separated by a 908 bp region containing no complete orfs. *pagO* encodes a putative protein of 304 amino acids (33.6 kDa) containing several hydrophobic regions, consistent with the prediction of PagO as an integral membrane protein. The N-terminus of this protein also contains a characteristic signal sequence.

To determine the probable transcriptional start of the *pagJ, K* and *M* genes, primer extension analysis was performed on RNA from $PhoP^-$ and $PhoP^c$ strains. As shown in Fig. 2, a product was observed in reactions containing $PhoP^c$ RNA but not in those with $PhoP^-$ RNA. This indicates that these genes are transcriptionally activated by PhoP-PhoQ as was predicted based on the regulation of Pag::PhoA fusion activity. This analysis shows that the *pagM* transcript probably begins 65 bases upstream of the ATG, with regions 5′ GCTAAC 3′ and 5′ TTGTGA 3′ functioning as the -10 and -35 sequences, respectively. Two products separated by 1 base are observed upon analysis of the *pagJ/K* transcript (as detailed below, *pagJ* and *pagK* are nearly identical in sequence). These may represent different transcriptional start sites of *pagJ* and *pagK*, or they may show that transcription of one or both genes can begin at either site. The *pagJ/K* transcriptional start begins 56 bp upstream of the ATG, with regions 5′ TTTAAT 3′ and 5′ TCACCT 3′ likely functioning as the -10 and -35 sequences, respectively.

Comparison of the *pagK/J* and *M* promoter regions to those of *pagC* [24], *pagD* [27] and *phoP* [19], which have been previously defined, failed to reveal any common DNA regions likely to form a PhoP-binding site. In addition, no inverted or direct repeats, which commonly form binding sites of regulatory proteins, are evident.

## *pagJ* and *pagK* are located within duplicated regions of the chromosome and are unique to *S. typhimurium* and *S. enteritidis*

In addition to the *pagK, M* and *O* linkage uncovered by comparisons between the *pag* DNA sequences, this analysis also showed identity between *pagJ* and *pagK*. Chromosomal mapping of these loci demonstrated that they were not linked. By a combination of pulsed-field gel electrophoresis and dot blot hybridization to a set of DNA from induced P22 Mu*d* prophages containing defined pieces of bacterial DNA [40], *pagJ* was shown to be located at 23–25 centisomes and *pagK* at 40–42 centisomes. *pagJ* and *pagK* show 90% identity at the DNA level and 83% identity of the predicted protein products. Sequence analysis of regions extending downstream and upstream of *pagJ* and *pagK* also show identity (95%). The duplicated region spans 1688 nucleotides, with endpoints 570 bp from the 3′ end of *pagK/J* to 920 bp from the 5′ end of *pagJ/*

**Table 1.** Database similarities to *pag* loci DNA and predicted amino acid sequence

| Gene/ region | Similarity[a] | Protein identity % | Protein similarity % | DNA identity % |
|---|---|---|---|---|
| *pagA*[b] | UDP-glucose dehydrogenase (orf in *Salmonella enterica cld* 5′ region (1–109) | 100 | 100 | 100 |
| | *E. coli* UDP-glucose dehydrogenase (1–388) | 88 | 95 | 78 |
| | *Vibrio cholerae* UDP-glucose dehydrogenase (1–388) | 74 | 89 | 67 |
| | *Streptococcus pneumoniae* UDP-glucose dehydrogenase (1–394) | 65 | 81 | 64 |
| *pagH* | *E. coli* Ais (24–107) | 77 | 85 | 68 |
| *pagN*[b] | *S. typhimurium orf* downstream of *sinR* (1–89) | 100 | 100 | 100 |
| | *E. coli* Tia ( 1–251) | 40 | 49 | 37 |
| | *S. typhimurium* PagC (100–164) | 34 | 49 | 49 |
| *pagO* | *Buchnera aphidicola* orf (1–290) | 55 | 72 | 58 |
| | *yadA* downstream orf of *Yersinia enterocolitica* (1–164) | 56 | 75 | 60 |
| | 5′ region of the *cat*1 gene of *Clostridium kluyveri* (1–131) | 20 | 37 | 41 |
| *pagP*[c] | *E. coli* CrcA (22–81) | 73 | 83 | 75 |
| *pagJ* region | | | | |
| 60–222 | tip of tail fibre-T4 (130–197) | 57 | 76 | 61 |
| 234–776 | tail fibre assembly-*E. coli* (1–179) | 65 | 82 | 66 |
| | tail fibre assembly-lambda phage (1–183) | 61 | 77 | 66 |
| 246–740 | *E. coli* and *Shigella boydii pin* 5′ (11–194) | 31 | 52 | 52 |
| 489–803 | *S. typhimurium nanH* 3′ (1–105) | 66 | 82 | 76 |
| 1465–1674 | IS*3* transposase (37–73, 274–288) | 69 | 87 | 78 |
| 1684–1803 | *E. coli* and *Y. enterocolitica* ArsB (108–140, 190–201) | 69 | 84 | 78 |
| 2051–2101 | *E. coli pin* 5′ (254–270) | 47 | 70 | 49 |
| 2053–3012 | *E. coli rhsD* (bp 6886–7394) | NA[d] | NA | 79 |
| 2131–3164 | *Actinobacillus* orf (1–346) | 73 | 84 | 71 |
| *pagK*MO region | | | | |
| 292–763 | tail fibre assembly-*E. coli* (22–179) | 55 | 74 | 56 |
| | tail fibre assembly-lambda phage (25–183) | 52 | 68 | 61 |
| 389–565 | *E. coli pin* 5′ (234–243) | 44 | 55 | 52 |
| 623–790 | *pin* 5′ region- *S. boydii* (119–174) | 51 | 66 | 52 |
| 476–790 | *S. typhimurium nanH* 3′ (1–105) | 67 | 75 | 80 |
| 1452–1681 | IS*3* transposase (32–79, 274–288) | 65 | 80 | 76 |

*continued*

**Table 1.** Database similarities to *pag* loci DNA and predicted amino acid sequence—*cont'd*

| Gene/ region | Similarity[a] | Protein identity % | Protein similarity % | DNA identity % |
|---|---|---|---|---|
| 1663–2126 | *E. coli* and *Y. enterocolitica arsB* (105–140, 190–201) | 65 | 81 | 73 |
| 2480–2587 | DNA invertases (139–174) | 83 | 88 | 71 |
| 2441–2470 | *E. coli pin* 5′ (234–243) | 60 | 90 | 62 |

[a] The amino acid residues (or DNA bases for *E. coli rhsD*) exhibiting similarity are noted in parentheses.
[b] Because the portion of sequenced DNA was 100% identical to a *Salmonella* gene currently in the database, the entire gene sequence from the database was used in subsequent comparisons. An exception is for similarity between PagN and PagC, where data from only the most similar region is presented.
[c] The *pagP* region sequenced was similar to the 5′ end of the *crcA* gene, part of which encodes the signal sequence. Because the regions encoding the signal sequences of these genes showed less identity than the rest of the sequence, only the predicted coding sequence for the mature protein was used in comparisons.
[d] NA, not applicable.

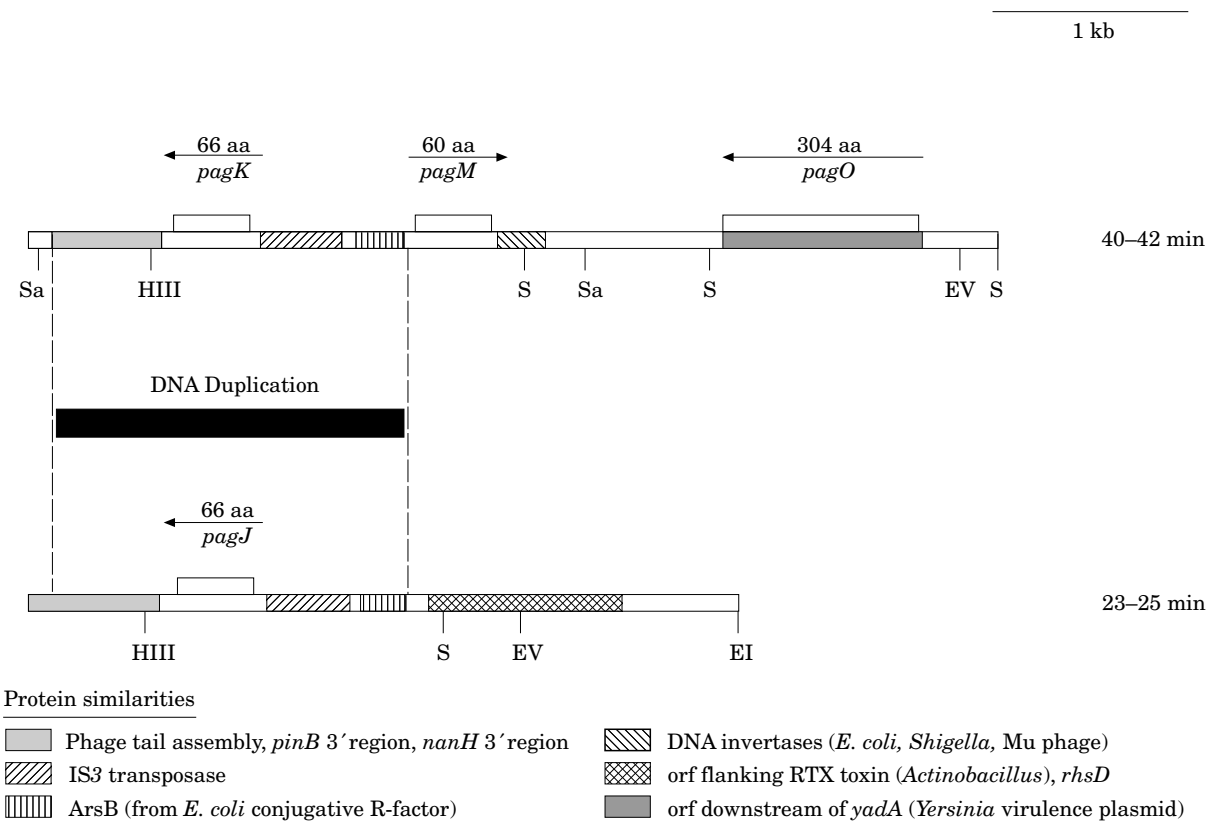


**Figure 1.** Partial restriction map of the *pagKMO* and *pagJ* regions showing the DNA duplication and regions that exhibit similarity to DNA or proteins in the database. Chromosomal map locations are shown to the right. Arrows denote the direction of transcription. Restriction enzyme abbreviations: EI, *Eco*RI; EV, *Eco*RV; HIII, *Hind*III; S, *Sac*I; SA, *Sal*I.

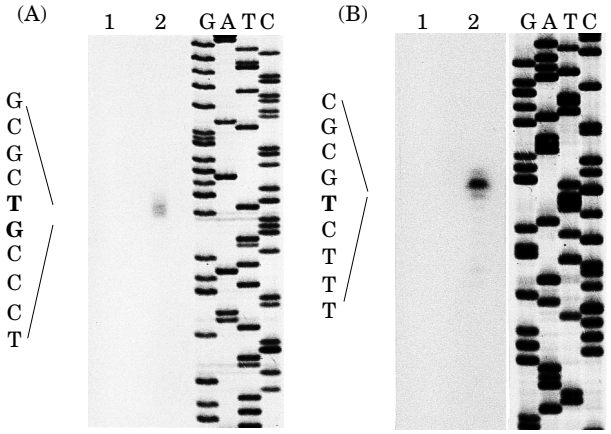


**Figure 2.** Primer extension analysis of *pagJ/K* (A) and *pagM* (B). Lane 1, the result of primer extension analysis with PhoP$^-$ RNA. Lane 2, the result of primer extension analysis with PhoP$^c$ RNA. Lanes GATC are DNA sequencing reactions initiated with the oligos used in the primer extension analysis. To the left of each panel is the sequence surrounding the predicted transcriptional start site (which is shown in bold).

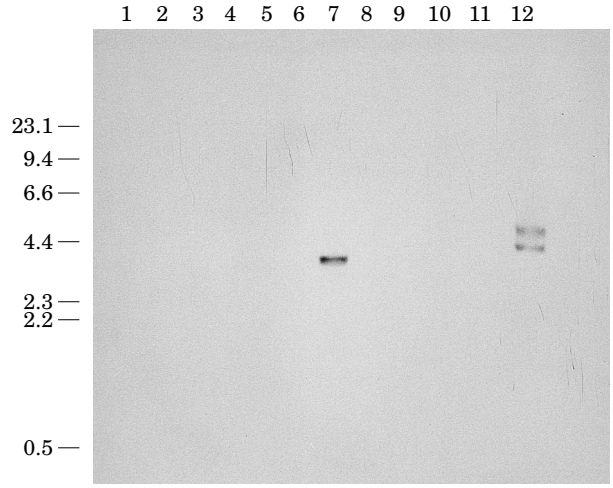


**Figure 3.** High stringency Southern blot hybridization of DNA digested with *Eco*RI from selected Enterobacteriaceae to a *pagK* specific probe . Lanes: 1, *E. coli* clinical isolate; 2, *Citrobacter freundii*; 3, *Klebsiella pneumoniae*; 4, *Shigella flexneri*; 5, *Yersinia enterocolitica*; 6, *Vibrio cholerae*; 7, *Salmonella enteriditis*; 8, *Salmonella paratyphi* A; 9, *Salmonella paratyphi* C; 10, *E. coli* DH5$\alpha$; 11, *Salmonella typhi* Ty-2; 12, *Salmonella typhimurium*. Molecular size markers are in kilobases.

*K*, which in the *pagKMO* region, lies 16 bp into the *pagM* orf (Fig. 1). Therefore, the duplicated region contains *pagK* and *pagJ*, but not *pagM* or *pagO*. The endpoints of this duplication do not contain any DNA repeats or regions that could form strong stem loops.

To examine the distribution of *pagJ/K* among other *Salmonella* serotypes and other members of the Enterobacteriaceae, Southern blot analysis was performed with a *pagJ/K*-specific probe. Figure 3 shows that under high stringency conditions, hybridizing fragments were observed in *S. typhimurium* (two bands representing *pagJ* and *pagK*) and *S. enteritidis* (one band), but no other hybridizing fragments were observed with the DNA of the human-specific *Salmonella* serotypes *S. typhi* or *S. paratyphi*, or to other Enterobacteriaceae tested. Several other *S. typhimurium* isolates were also analysed by Southern blot, and these showed identical hybridization patterns to that seen with *S. typhimurium* 14028s (data not shown). Similar experiments using a probe outside the duplication (covering most of the *pagM* gene) showed hybridization to all *Salmonella* serotypes tested and *Klebsiella pneumonia* (data not shown).

## Translation of sequences surrounding *pagJ* and *pagK* show similarity to proteins associated with mobile or extrachromosomal DNA

Comparison of the DNA and the six-frame translation of the *pagJ* region and *pagKMO* region to the database resulted in numerous high-scoring similarities; however, no similarities were found for *pagJ, K* or *M* at the DNA level or with the predicted protein translation. The predicted product of the *pagO* gene, as mentioned previously, was similar to 3 orfs of unknown function. Several other partial similarities were observed in intergenic regions. These similarities were of 35 or more amino acids and do not correspond to known functional domains of the proteins showing similarity. The regions of similarity are shown graphically in Fig. 1 and the details of these similarities are described in Table 1. Downstream of the *pagK* and *pagJ* genes is a region with similarity to bacteriophage tail fibre genes [41, 42], a segment of the *E. coli* and *Shigella pinB* 5′ region [43, 44], and the *S. typhimurium nanH* 3′ region [45]. The *pagJ* or *pagKMO* regions are not, however, located near *nanH*. The endpoint of the duplication is within the homology to the tail fibre assembly genes, and this homology stops at the point of the duplication in the *pagKMO* region, but continues past the duplication point in the *pagJ* region.

Between the *pagK* and *M* genes, located within the duplication, are similarities to IS*3* transposase [46] and to the *arsB* protein, which is involved in resistance to arsenic and encoded on plasmids of *Y. enterocolitica* and *E. coli* [47,

48]. The ArsB similarity is in the same frame and continues six amino acids into the predicted *pagM* protein, stopping at the endpoint of the duplication. Downstream of the *pagM* gene are regions similar to phage and bacterial DNA invertases and to the *pin* 5′ region of *E. coli* (a different region from what is similar to the region downstream of *pagK*). Upstream of *pagJ*, past the duplication, is a segment with similarity to the *E. coli rhsD* (*r*ecombination *h*ot *s*pot) gene region [49] and an *Actinobacillus pleuropneumoniae* orf that is present in two copies, both of which flank a cytotoxin [50]. These *Actinobacillus* orfs can recombine at a low frequency to delete the cytotoxin. This *Salmonella* region, although similar to the entire *A. pleuropneumoniae* orf, contains several frameshifts and is unlikely to produce a functional protein.

The *pagJ* and *pagKMO* regions contain a G+C ratio lower than that of the entire Salmonella genome (*pagJ* region, 45%; *pagJ*, 41%; *pagKMO* region, 43%; *pagK*, 37%; *pagM*, 48%; *pagO*, 41%; *Salmonella* genome, 52%). Because these regions contain a G+C ratio lower than that of the *Salmonella* genome average, it suggests that these DNA segments may have been acquired by horizontal transmission.

The partial similarities in the *pagJ* and *pagKMO* regions correspond to segments of known plasmids, transposons, phage, and invertable DNA elements. Therefore, this region may represent/have represented a location for multiple DNA recombination events, and the observed similarities are remnants of the genes that were carried by or involved in DNA movement into or out of this chromosomal region.

### Deletions of *pagJ*, *K* and *M* do not attenuate mouse virulence

*TnphoA* insertions in *pagJ, pagK* or *pagM* have previously been shown to result in a virulence defect in the BALB/c mouse model [23]. Because we have previously seen attenuation of virulence due to insertion of Tn*phoA* in *pagC* or *pagD* that returned to wild type virulence when deletion mutants were constructed [27] (S.I. Miller & J. Williams, unpublished data), we created individual and combinatorial deletions of *pagJ, pagK* and *pagM*. Deletions of these three genes individually or in any combination did not attenuate strain virulence (data not shown). Deletions of these three genes combined with deletions of *pagC* and *pagD* also did not alter *S. typhimurium* virulence.

## Discussion

PhoP-PhoQ activate transcription of a number of *S. typhimurium* genes within the infected host in response to environmental signals. Based on the phenotype of a PhoP-PhoQ null strain, these genes probably encode products that contribute to virulence, and specifically, survival within host tissues. We have previously described the identification of a number of *pag*, and here we present the cloning, sequencing, and partial characterization of PhoP-activated loci.

DNA sequence was determined from clones of the *pag*-transposon junctions, and upon comparison of these sequences to those in the database, several similarities (DNA and protein) were observed. PagA (PmrE) is similar to UDP-glucose dehydrogenases, which are involved in capsule or complex carbohydrate synthesis by a number of bacterial species [28-31]. This enzyme is necessary for the addition of aminoarabinose to lipid A (probably in the biosynthesis of aminoarabinose) and for resistance to the antimicrobial compound, polymyxin. *pagA (pmrE)* is regulated by PmrA-PmrB, which is itself regulated by PhoP-PhoQ [20, 21]. We and others have shown this gene to be regulated by acid, which is probably due to the fact that mild acidification is a signal for induction of the PmrA-PmrB system [12, 21, 32] (J.S. Gunn, unpublished data). The gene located downstream of *pagA* is *cld*, which is involved in determining the length of the LPS oligosaccharide chain [51]. Although the PhoP-PhoQ and PmrA-PmrB systems both regulate genes that modify the structure of *Salmonella* LPS including chain length [9], the *cld* gene is not co-transcribed with *pagA* nor is it regulated by PhoP-PhoQ when grown under laboratory conditions [33].

The predicted *pagH (pmrG)* product is directly regulated by PmrA-PmrB and not PhoP-PhoQ [33], and is similar to the aluminum-inducible Ais protein of *E. coli* [36]. The gene encoding the Ais protein was discovered by screening a bank of random Tn*5-luxAB* fusions for induction in the presence of various metals. It is not known if the induction of *ais* transcription is due to a direct or an indirect effect of aluminium. Preliminary experiments suggest that PhoP-PhoQ

is not responsive to aluminium (J. Gunn, unpublished results), but it is not known whether the PmrA-PmrB regulon can be activated or repressed by aluminium.

*pagN* is located downstream of the *S. typhimurium sinR* gene, located in a chromosomal region shown to be confined to the Salmonellae [34]. This gene was also identified as an *in vivo* induced gene by IVET [26]. These authors noted that a gene upstream of *pagN* was also regulated by PhoP-PhoQ and, therefore, these genes may be co-transcribed. The predicted PagN protein sequence is similar to a family of outer membrane proteins that include Tia of enteropathogenic *E. coli*, Ail of *Yersinia* spp. and PagC. Direct comparison of PagN and PagC revealed regions of high identity, including 49% similarity over an internal stretch of 68 amino acids (data not shown). Therefore, it is possible that *pagC* and *pagN* encode proteins of similar function. However, overall, these proteins are less similar than PagC and Rck, a non-PhoP regulated *S. typhimurium* member of this family of proteins [52].

The *pagP* gene product is homologous to the *E. coli* CrcA protein, which is one of three linked genes (*crcA, cspE*, and *crcA*) that are required for resistance to camphor [38]. It has been demonstrated that exposure to camphor results in chromosome decondensation, and that overexpression of these genes promotes or protects chromosome folding. However, it is also possible that resistance to camphor involves a change in the cell surface that effectively excludes camphor and results in spontaneous chromosome recondensation. In fact, preliminary data in our laboratory suggests that *pagP* is involved in LPS modifications and resistance of *Salmonella* to antimicrobial peptides [53].

Three PhoP-activated genes, *pagK, M* and *O*, were shown to be linked on the chromosome. In addition, *pagK* and *pagJ* were nearly identical in DNA and protein sequence, as these two genes each reside within a 1.6 kb duplicated region of the chromosome. Analysis of DNA within and immediately outside this duplication revealed a strong similarity to a variety of proteins, including transposases, DNA invertases, phage proteins, and proteins encoded on plasmids. The majority of these segments are non-coding and do not correspond to functional domains of the proteins that exhibit similarity. Besides containing what appears to be gene fragments left behind from previous recombination events, this region has a G+C content considerably lower than the average G+C content of the *S. typhimurium* genome. Also, the *pagJ* and *pagK* genes, which are located within the duplication, are found only in *Salmonella* serotypes Typhimurium and Enteritidis, and not in any of the other Enterobacteriaceae tested. The two serotypes containing *pagJ* and *pagK* have a broad host range, while those *Salmonella* serotypes not hybridizing to the *pagK* specific probe are strictly human pathogens, suggesting a possible correlation of these genes and host range. Therefore, it appears that this is/was a dynamic region of the chromosome in which sections were obtained by Salmonella through inter-species transmission. It is interesting to speculate that mobility of a functional transposon or phage may have been responsible for the duplication of *pagJ* and *K*. Preliminary sequence analysis of DNA outside of the reported *pagJ* region show a segment exhibiting both DNA and protein similarity to a target of the *Shigella* and *Yersinia* spp. TypeIII secretion systems, further demonstrating the likelihood that multiple horizontal transmission events took place in these chromosomal regions (E. A. Miao & S. I. Miller, unpublished data).

In addition to *pagJ* and *pagK*, PhoP-activated genes *pagC, pagD, mgtAB* and *phoN* are located in regions of low G+C content and are thought to have been acquired by horizontal transmission [15, 18, 27, 54]. Like *pagJ* and *pagK*, several PhoP-activated genes have been shown by DNA hybridization experiments to be found in *Salmonella* but absent in the DNA of other enteric bacteria. Therefore, from an evolutionary standpoint, it is interesting that PhoP-PhoQ has come to regulate several acquired genes which are likely to function together to increase *Salmonella* survival within host tissues and macrophage phagosomes.

$LD_{50}$ analysis of strains carrying *pag*::Tn*phoA* fusions showed that disruption of the *pagC, pagD, pagJ, pagK* and *pagM* loci resulted in attenuation of virulence [1, 23]. However, in a surprising finding, strains carrying single deletion mutations, or combinations of deletions of these genes, did not result in an increase in $LD_{50}$. It is possible that the observed virulence defect for the *pag*::Tn*phoA* strains could be merely an artefact of overexpression of the Pag-PhoA fusion protein. This seems unlikely because many other highly expressed Pag-PhoA fusions do not result in virulence attenuation. One potential explanation for these results is that the Pag-PhoA fusion exhibits a dominant-negative effect on virulence through interaction

with a virulence factor(s), which may include other Pag. Upon deletion of the *pag* loci, a similar protein may substitute for this gene product to result in a functional interaction with the virulence factor. In support of a protein interaction and not an overexpression theory, conservative amino acid changes in the PagC KYRYE sequence (found in all members of the PagC-family of proteins excluding PagN) resulted in attenuation of *S. typhimurium* virulence (S. I. Miller & J. Williams, unpublished data).

We have presented data here that suggests functional redundancy of *pag* gene products within the PhoP-PhoQ regulon. PagJ and PagK are nearly identical in protein sequence, and probably perform identical functions. *pagD, J, K* and *M* are predicted to encode mature proteins less than 8.5 kDa, and it is possible that these, and potentially other yet undefined *pag*, comprise a family of small molecular weight envelope or secreted proteins with similar function. Data was also presented concerning the similarity between the predicted outer membrane proteins PagN and PagC. In addition, the PhoP-activated magnesium transporters *mgtA* and *mgtBC* also encode proteins of redundant function, as both encode high affinity $Mg^{2+}$ uptake systems [15, 55]. Recently, it was demonstrated that individual deletions of these transport systems had little or no effect on virulence; however, deletion of both systems resulted in a significant decrease in mouse virulence [18]. Lastly, it has been shown that PhoP-PhoQ regulate at least three independent mechanisms of resistance to various types of antimicrobial peptides [53]. Therefore, to date, no single PhoP-activated gene mutation results in the identical phenotype or magnitude of a phenotype attributed to a PhoP-null strain. The data outlined above suggest that functional redundancies within the regulon are a potential explanation for this phenomenon.

Information concerning Pag and the role of these proteins as members of a global regulon central to Salmonellae virulence is increasing. However, it is likely that additional functional and genetic redundancies will be uncovered upon the identification and characterization of more *pag*. This level of complexity will make it difficult to identify those Pag essential for virulence. Therefore, this work represents an initial step at obtaining a more complete understanding of the genes regulated by PhoP-PhoQ, which is necessary to determine the molecular mechanisms of this regulon essential for Salmonellae pathogenesis.

## Materials and methods

### Bacterial strains and growth conditions

Bacterial strains and plasmids used in this study are described in Table 2, with the exception of the strains containing transposon insertions defining *pagE-P* [23] and those used for Southern blot analysis [27]. Cultures were grown at 37°C with aeration in Luria Broth [56]. Antibiotics were used at the following concentrations: chloramphenicol, 25 μg/ml; kanamycin, 45 μg/ml; tetracycline, 8–25 μg/ml; streptomycin, 1 mg/ml, and ampicillin, 50 μg/ml.

### Bacterial genetics

Chromosomal DNA was isolated by the method described by Mekalanos *et al*. [57]. Southern blotting and detection of hybridization was accomplished as previously described [27]. P22HT-*int* bacteriophage was used in all transductional crosses [58,59] and all strains resulting from transductional crosses were routinely confirmed to be non-lysogens of P22. Mapping of cloned DNA was accomplished by hybridization of labelled probes to blots of digested DNA separated by pulsed-field gel electrophoresis, or by hybridization to a dot blot of DNA from a set of random, induced P-22 prophages. When induced, these prophages generate transducing lysates that contain bacterial DNA flanking one side of the insertion [40].

### Plasmid construction and cloning of *pag* DNA

*pag* sequences were cloned by one of two methods. One method involved the construction of genomic libraries from *pag*::Tn*phoA* or *pag*::Mu*d*J strains, followed by *E. coli* transformation and screening for kanamycin (encoded by the transposon) and ampicillin (encoded by the plasmid) resistances. The *pagA*::Mu*d*J 3′ fusion junction was cloned from a *Sac*I library in pBCSK+, and *pagF* was cloned from a *Eco*RV-*Sal*I library in pBluescriptSK+. A clone of DNA 3′ to the *pagJ*::Tn*phoA* transposon insertion was also identified in this manner. This clone was identified by digesting the genomic DNA with *Eco*RI (in Tn*phoA*) and *Eco*RV and ligating to pBluescriptSK+. The second method of cloning *pag* sequences involved the construction and use

**Table 2.** Bacterial strains, plasmids and relevant properties

| Strains | Genotype/relevant phenotype | Source or reference |
|---|---|---|
| *Salmonella* | | |
| CS019 | *phoN2 zxx*::6251 Tn*10*d-cam | [1] |
| ATCC14028s | wildtype | ATCC |
| CS012 | ATCC14028s *pagA*::Mu*d*J | [1] |
| JSG540 | *pagC-D, K-M* deletion | This work |
| JSG543 | *pagC-D, K-M, J* deletion | This work |
| JSG542 | CS019 streptomycin resistant | This work |
| *E. coli* | | |
| DH5α | F-ø 80d*lac*ZΔM15 Δ(*lacZYA-argF*) U169*endA1recA*1*hsdR*17deoR *thi*-1*supE*44l-*gyrA*96*relA*1 | BRL |
| SM10λPir | *thi*-1 *thr*-1 *leuB*6 *supE*44 *tonA*21*lacY*1*recA*::RP4-2-Tc::Mu | |
| *plasmids* | | |
| pSG20 | pBCKS+ with a 16 kb *Sac*I fragment containing the *pagA*:: Mu*d*J 3′ fusion junction | This work |
| pBB01 | Pir-dependant suicide vector used to clone *pag*::Tn*phoA* 5′ fusion junctions | This work |
| pBB14EE | pBB01 containing the *pagE*:: Tn*phoA* 5′ fusion junction (*Eco*RV) | This work |
| pBKF | pBluescript IISK+ with an *Eco*RV-*Sal*I fragment containing the 5′ *pagF*:: Tn*phoA* fusion junction | This work |
| pBB03SG | pBB01 containing the *pagG*:: Tn*phoA* 5′ fusion junction (*Sal*I) | This work |
| pBB01SH | pBB01 containing the *pagH*:: Tn*phoA* 5′ fusion junction (*Sal*I) | This work |
| pBB13BI | pBB01 containing the *pagI*:: Tn*phoA* 5′ fusion junction (*Bgl*II) | This work |
| pBB05BJ | pBB01 containing the *pagJ*:: Tn*phoA* 5′ fusion junction (*Bgl*II) | This work |
| pBB02SK | pBB01 containing the *pagK*:: Tn*phoA* 5′ fusion junction (*Sal*I) | This work |
| pBB04EL | pBB01 containing the *pag*:: Tn*phoA* 5′ fusion junction (*Eco*RV) | This work |
| pBB15KM | pBB01 containing the *pagM*:: Tn*phoA* 5′ fusion junction (*Kpn*I) | This work |

*continued*

**Table 2.** Bacterial strains, plasmids and relevant properties—*cont'd*

| Strains | Genotype/relevant phenotype | Source or reference |
|---|---|---|
| pBB09BN | pBB01 containing the *pagN*::Tn*phoA* 5′ fusion junction (*Bgl*II) | This work |
| pBB16EO | pBB01 containing the *pagO*::Tn*phoA* 5′ fusion junction (*Eco*RV) | This work |
| pBB12SP | pBB01 containing the *pagP*::Tn*phoA* 5′ fusion junction (*Sal*I) | This work |
| pKMO1 | pUC19 with a *Hin*dIII-*Eco*RI fragment containing the *pagKMO* genes | This work |
| pBJS1 | PCR generated clone containing the *pagJ* gene | This work |
| pJTK1 | pBluescript IISK+ with an *Eco*RI-*Eco*RV fragment containing sequences downstream of *pagJ* | This work |
| pCVD442 | positive selection (*sacB*) suicide vector | [64] |
| pKAS32 | positive selection (*rpsL*) suicide vector | [65] |

ATCC, American Type Culture Collection; BRL, Bethesda Research Laboratory.

of the vector pBB01. pBB01 is pGP704, a pir-dependent suicide vector, containing a PCR generated fragment of the *phoA* gene cloned into the *Sma*I-*Eco*RI sites. Upon recombination of pBB01 onto the chromosome of a *pag*::Tn*phoA* strain, chromosomal DNA was isolated from the recombinant and digested with an enzyme located in the pBB01 multiple cloning site. Digested DNA was self-ligated and transformed into a Pir-producing *E. coli* host. This procedure effectively cloned DNA 5′ to the Tn*phoA* insertion. A wildtype (wt) clone of the *pagJ* gene, pBJS1, obtained by PCR, was used to confirm the sequence at the transposon insertion site. A pUC19 *Hin*dIII-*Eco*RI clone from wt DNA encompassing the *pagK, M* and *O* genes, pKMO1, was identified by hybridization to a *pagK* specific probe.

## DNA sequence analysis

DNA sequence analysis was performed on double-stranded templates by the dideoxy-chain termination method [60] using the Sequenase enzyme and kit reagents (US Biochemicals Corp., Cleveland, OH, U.S.A.). Computer analysis of the DNA sequence was accomplished with the Wisconsin package (GCG, version 7) program.

## Construction of chromosomal deletions

Deletion of the *pagK-M* and *J* loci were accomplished as follows. A *pagK-M* double deletion was constructed by first amplifying regions downstream of *pagK* and upstream of *pagM* by PCR. The *pagK* downstream fragment contained a *Eco*RV site at the far end and an *Nco*I site at the end near *pagK*, which were incorporated as part of the primers used in the amplification. The *pagM* fragment contained an *Nco*I site at the end near *pagM* and an *Xba*I site at the distal end. These fragments were each digested with *Nco*I, ligated, digested with *Xba*I and *Eco*RV, and ligated into pBluescriptSK+. The insert from this clone containing the *pagK-M* deletion was ligated into the *Sma*I and *Xba*I sites of the suicide vector pCVD442 [61]. Upon

recombination onto the chromosome, subsequent selection on sucrose plates, and Southern hybridization analysis of strains with potential deletions, a strain containing the defined *pagK-M* double deletion was identified.

The *pagJ* deletion was constructed in a manner similar to the *pagK-M* deletion, with one PCR fragment having *Kpn*I and *Xho*I restriction sites at its ends, and the other fragment with *Xba*I and *Xho*I sites at its ends. These fragments were digested with *Xho*I and ligated. Following ligation, these fragments were digested with *Xba*I and *Kpn*I and ligated to the vector pKAS32 [62], which contains the *rpsL* gene for positive selection. *E. coli* containing the *pagJ* deletion in pKAS32 were mated with a streptomycin resistant strain to recombine the deletion into the chromosome. This strain was then subjected to selection on plates containing streptomycin to identify strains that deleted the plasmid sequences and the *rpsL* gene. Southern hybridization analysis of strains with potential deletions identified a strain containing the defined *pagJ* deletion.

### Promoter analysis

RNA was isolated from *Salmonella* strains by the hot phenol procedure [63]. Primers (0.2 pmol) end-labelled with [$\gamma$-$^{32}$P]ATP were annealed to *S. typhimurium* RNA (10 μg) and extended with reverse transcriptase. Reactions were electrophoresed adjacent to sequencing reactions initiated with the primers used for cDNA synthesis.

### Nucleotide sequence accession number

The DNA sequences described have been deposited in GenBank and assigned accession numbers AF013775 (*pagK,M,O* region) and AF013776 (*pagJ* region).

## Acknowledgements

We thank members of the laboratory for helpful discussions and comments on this manuscript. Special thanks to Dean Thompson for his technical contributions. This work was supported by a grant from the NIH to SIM (AI 30479).

## References

1 Miller SI, Kukral AM, Mekalanos JJ. A two-component regulatory system (*phoP phoQ*) controls *Salmonella typhimurium* virulence. *Proc Natl Acad Sci USA* 1989; **86**: 5054–8.
2 Groisman EA, Chiao E, Lipps CJ, Heffron F. *Salmonella typhimurium phoP* virulence gene is a transcriptional regulator. *PNAS USA* 1989; **86**: 7077–81.
3 Fields PI, Swanson RV, Haidaris CG, Heffron F. Mutants of *Salmonella typhimurium* that cannot survive within the macrophage are avirulent. *Proc Natl Acad Sci USA* 1986; **83**: 5189–93.
4 Hohmann EL, Oletta CA, Killeen KP, Miller SI. *phoP/phoQ*-deleted *Salmonella typhi* (TY800) is a safe and immunogenic single dose typhoid fever vaccine in volunteers. *J Infect Dis* 1996; **173**: 1408–14.
5 Fields PI, Groisman EA, Heffron F. A *Salmonella* locus that controls resistance to microbicidal proteins from phagocytic cells. *Science* 1989; **243**: 1059–62.
6 Groisman EA, Parra-Lopez C, Salcedo M, Lipps CJ, Heffron F. Resistance to host antimicrobial peptides is necessary for *Salmonella* virulence. *Proc Natl Acad Sci USA* 1992; **89**: 11939–43.
7 Miller SI, Pulkkinen WS, Selsted ME, Mekalanos JJ. Characterization of defensin resistance phenotypes associated with mutations in the *phoP* virulence regulon of *Salmonella typhimurium. Infect Immun* 1990; **58**: 3706–10.
8 Wick MJ, Harding CV, Twesten NJ, Normark SJ, Pfeifer JD. The *phoP* locus influences processing and presentation of *S. typhimurium* antigens by activated macrophages. *Mol Microbiol* 1995; **16**: 465–76.
9 Guo L, Lim K, Gunn JS, *et al*. Regulation of lipid A modifications by *Salmonella typhimurium* virulence genes *phoP-phoQ. Science* 1997; **276**: 250–53.
10 Gunn JS, Hohmann EL, Miller SI. Transcriptional regulation of *Salmonella* virulence: a PhoQ periplasmic domain mutation results in increased net phosphotransfer to PhoP. *J Bact* 1996; **178**: 6369–73.
11 Miller SI, Mekalanos JJ. Constitutive expression of the PhoP regulon attenuates *Salmonella* virulence and survival within macrophages. *J Bact* 1990; **172**: 2485–90.
12 Alpuche-Aranda CM, Swanson JA, Loomis WP, Miller SI. *Salmonella typhimurium* activates virulence gene transcription within acidified macrophage phagosomes. *Proc Natl Acad Sci USA* 1992; **89**: 10079–83.
13 Waldburger CD, Sauer RT. Signal detection by the PhoQ sensor-transmitter. *J Biol Chem* 1996; **271**: 26630–6.
14 Garcia-Vescovi E, Ayala YM, DiCera E, Groisman EA. Characterization of the bacterial sensor protein PhoQ. Evidence for distinct binding sites for $Mg^{2+}$ and $Ca^{2+}$. *J Biol Chem* 1997; **272**: 1440–3.
15 Garcia-Vescovi E, Soncini FC, Groisman EA. $Mg^{2+}$ as an extracellular signal: environmental regulation of Salmonella virulence. *Cell* 1996; **84**: 165–74.
16 Hueck CJ, Hantman MJ, Bajaj V, *et al*. *Salmonella typhimurium* secreted invasion determinants are homologous to *Shigella* Ipa proteins. *Mol Microbiol* 1995; **18**: 479–90.
17 Pegues DA, Hantman MJ, Behlau I, Miller SI. PhoP/PhoQ transcriptional repression of *Salmonella typhimurium* invasion genes: evidence for a role in protein secretion. *Mol Microbiol* 1995; **17**: 169–81.
18 Blanc-Potard AB, Groisman EA. The *Salmonella selC*

locus contains a pathogenicity island mediating intramacrophage survival. *EMBO J* 1997; **16**: 5376–85.
19 Soncini FC, García Véscovi E, Groisman EA. Transcriptional autoregulation of the *Salmonella typhimurium phoPQ* operon. *J Bact* 1995; **177**: 4364–71.
20 Gunn JS, Miller SI. PhoP/PhoQ activates transcription of *pmrA/B*, encoding a two-component system involved in *Salmonella typhimurium* antimicrobial peptide resistance. *J Bact* 1996; **178**: 6857–64.
21 Soncini FC, Groisman EA. Two-component regulatory systems can interact to process multiple environmental signals. *J Bact* 1996; **178**: 6796–801.
22 Kier LD, Weppelman R, Ames BN. Regulation of two phosphatases and a cyclic phosphodiesterase of *Salmonella typhimurium. J Bact* 1977; **130**: 420–8.
23 Belden WJ, Miller SI. Further characterization of the PhoP regulon: identification of new PhoP-activated virulence loci. *Infect Immun* 1994; **62**: 5095–101.
24 Pulkkinen WS, Miller SI. A *Salmonella typhimurium* virulence protein is similar to a *Yersinia enterocolitica* invasion protein and a bacteriophage lambda outer membrane protein. *J Bact* 1991; **173**: 86–93.
25 Soncini FC, Garcia Vescovi E, Solomon F, Groisman EA. Molecular basis of the magnesium deprivation response in *Salmonella typhimurium*: identification of PhoP-regulated genes. *J Bact* 1996; **178**: 5092–9.
26 Heithoff DM, Conner CP, Hanna PC, *et al*. Bacterial infection as assessed by *in vivo* gene expression. *Proc Natl Acad Sci USA* 1997; **94**: 934–39.
27 Gunn JS, Alpuche-Aranda CA, Loomis WP, Miller SI. Characterization of the *S. typhimurium pagC/pagD* chromosomal region. *J Bact* 1995; **177**: 5040–7.
28 Dillard JP, Vandersea MW, Yother J. Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in Streptococcus pneumoniae. *J Exp Med* 1995; **181**: 973–83.
29 Drummelsmith J, Amor PA, Whitfield C. Polymorphism, duplication, and IS1-mediated rearrangement in the chromosomal his-rfb-gnd region of *Escherichia coli* strains withgroup IA and capsular K antigens. *J Bacteriol* 1997; **179**: 3232–38.
30 Comstock LE, Johnson JA, Michalski JM, Morris JGJ, Kaper JB. Cloning and sequence of a region encoding a surface polysaccharide of Vibrio cholerae O139 and characterization of the insertion site in the chromosome of *Vibrio cholerae* O1. *Mol Microbiol* 1996; **19**: 815–26.
31 Morona R, Mavris M, Fallarino A, Manning PA. Characterization of the rfc region of *Shigella flexneri. J Bacteriol* 1994; **176**: 733–47.
32 Valdivia RH, Falkow S. Bacterial genetics by flow cytometry: rapid isolation of Salmonella typhimurium acid-inducible promoters by differential fluorescence induction. *Mol Microbiol* 1996; **22**: 367–78.
33 Gunn JS, Lim KB, Krueger J, *et al*. PmrA-PmrB regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance. *Mol Microbiol* 1998; **27**: 1171–1182.
34 Groisman EA, Sturmoski MA, Solomon FR, Lin R, Ochman H. Molecular, functional, and evolutionary analysis of sequences specific to Salmonella. *Proc Natl Acad Sci USA* 1993; **90**: 1033–7.
35 Fleckenstein JM, Kopecko DJ, Warren RL, Elsinghorst EA. Molecular characterization of the tia invasion locus from enterotoxigenic *Escherichia coli. Infect Immun* 1996; **64**: 2256–65.
36 Guzzo A, DuBow MS. Identification and characterization of genetically programmed responses to toxic metal exposure in *Escherichia coli. FEMS Microbiol Rev* 1994; **14**: 369–74.
37 Skurnik M, Wolf-Watz H. Analysis of the *yopA* gene encoding the Yop1 virulence determinants of *Yersinia* spp. *Mol Microbiol* 1989; **3**: 517–29.
38 Hu KH, Liu E, Dean K, *et al*. Overproduction of three genes leads to camphor resistance and chromosome condensation in *Escherichia coli. Genetics* 1996; **143**: 1521–32.
39 Von Heinje G. Signal sequences:the limits of variation. *J Mol Biol* 1985; **184**: 99–105.
40 Benson NR, Goldman BS. Rapid mapping in *S. typhimurium* with Mu*d*-P22 Prophages. *J Bact* 1992; **174**: 1673–81.
41 Blattner FR, Plunkett III G, Bloch CA, *et al*. The complete genome sequence of *Escherichia coli* K-12. *Science* 1997; **277**: 1453–62.
42 Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB. Nucleotide sequence of bacteriophage lambda DNA. *J Mol Biol* 1982; **162**: 729–73.
43 Tominaga A, Ikemizu S, Enomoto M. Site-specific recombinase genes in three *Shigella* subgroups and nucleotide sequences of a *pinB* gene and an invertible B segmentfrom Shigella boydii. *J Bacteriol* 1991; **173**: 4079–87.
44 Plasterk RHA, van de Putte P. The invertible P-DNA segment in the chromosome of *Escherichia coli. EMBO J* 1985; **4**: 237–42.
45 Hoyer LL, Hamilton AC, Steenbergen SM, Vimr ER. Cloning, sequencing and distribution of the *Salmonella typhimurium* LT2 sialidase gene, *nanH*, provides evidence for interspecies gene transfer. *Mol Microbiol* 1992; **6**: 873–84.
46 Timmerman KP, Tu CP. Complete sequence of IS3. *Nucleic Acids Res* 1985; **13**: 2127–39.
47 Chen CM, Misra TK, Silver S, Rosen BP. Nucleotide sequence of the structural genes for an anion pump. The plasmid-encoded arsenical resistance operon. *J Biol Chem* 1986; **261**: 15030–8.
48 Neyt C, Iriarte M, Thi VH, Cornelis GR. Virulence and arsenic resistance in Yersiniae. *J Bact* 1997; **179**: 612–19.
49 Sadosky AB, Davidson A, Lin RJ, Hill CW. *rhs* gene family of *Escherichia coli* K-12. *J Bact* 1989; **171**: 636–42.
50 Anderson C, Potter AA, Gerlach GF. Isolation and molecular characterization of spontaneously occurring cytolysin-negative mutants of *Actinobacillus pleuropneumoniae* serotype 7. *Infect Immun* 1991; **59**: 4110–16.
51 Bastin DA, Stevenson G, Brown PK, Haase A, Reeves PR. Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid sysnthesis, with a novel mechanism for determining chain length. *Mol Microbiol* 1993; **7**: 725–34.
52 Heffernan EJ, Harwood J, Fierer J, Guiney D. The *Salmonella typhimurium* virulence plasmid complement resistance gene *rck* is homologous to a family of virulence-related outer membrane protein genes, including *pagC* and *ail. J Bact* 1992; **174**: 84–91.
53 Guo L, Lim KB, Poduje CM, *et al*. Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. 1998; Submitted for publication.
54 Groisman EA, Saier MH, Ochman H. Horizontal transfer of a phosphatase gene as evidence for mosaic structure of the *Salmonella* genome. *EMBO J* 1992; **11**: 1309–16.

55 Maguire ME. MgtA and MgtB: prokaryotic P-type ATPases that mediate Mg2+ influx. *J Bioenerg Biomembr* 1992; **24**: 319–28.
56 Miller JH. *Experiments in Molecular Genetics*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1972.
57 Mekalanos JJ. Duplication and amplification of toxin genes in *Vibrio cholerae. Cell* 1983; **35**: 253–63.
58 Chan RK, Botstein D, Watanabe T, Ogata Y. Specialized transduction of tetracycline resistance by phage P22 in *Salmonella typhimurium*. II. Properties of a high-frequency-transducing lysate. *Virology* 1972; **50**: 883–98.
59 Davis RW, Botstein D, Roth JR. *Advanced Bacterial Genetics: A Manual for Genetic Engineering*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1980.
60 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977; **74**: 5463–7.
61 Donnenberg MS, Kaper JB. Construction of an eae deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect Immun* 1991; **59**: 4310–17.
62 Skorupski K, Taylor RK. Positive selection vectors for allelic exchange. *Gene* 1996; **169**: 47–52.
63 Case CC, Roels S, Gonzales JE, Simons EL, Simons RW. Analysis of the promoters and transcripts involved in IS*10* anti-sense RNA control. *Gene* 1988; **72**: 219–36.