Transcription factor function and promoter architecture govern the evolution of bacterial regulons

J. Christian Perez[1] and Eduardo A. Groisman[2]

Department of Molecular Microbiology, Howard Hughes Medical Institute, Washington University School of Medicine, St. Louis, MO 63110

Edited by Jeffrey I. Gordon, Washington University School of Medicine, St. Louis, MO, and approved January 16, 2009 (received for review October 14, 2008)

Evolutionary changes in ancestral regulatory circuits can bring about phenotypic differences between related organisms. Studies of regulatory circuits in eukaryotes suggest that these modifications result primarily from changes in *cis*-regulatory elements (as opposed to alterations in the transcription factors that act upon these sequences). It is presently unclear how the evolution of gene regulatory circuits has proceeded in bacteria, given the rampant effects of horizontal gene transfer, which has significantly altered the composition of bacterial regulons. We now demonstrate that the evolution of the regulons governed by the regulatory protein PhoP in the related human pathogens *Salmonella enterica* and *Yersinia pestis* has entailed functional changes in the PhoP protein as well as in the architecture of PhoP-dependent promoters. These changes have resulted in orthologous PhoP proteins that differ both in their ability to promote transcription and in their role as virulence regulators. We posit that these changes allow bacterial transcription factors to incorporate newly acquired genes into ancestral regulatory circuits and yet retain control of the core members of a regulon.

cis-regulatory elements | horizontal gene transfer | PhoP | *Salmonella* | *Yersinia*

Gene regulatory networks undergo major modifications over time, and these modifications provide an important source of phenotypic diversity among closely related organisms (1, 2). In principle, these modifications may result from changes in *cis*-regulatory sequences and/or from alterations in the deployment and/or activity of the transcription factors (TFs) that act on these sequences. In eukaryotes, in which most studies of the evolution of gene regulation have been carried out, most regulatory changes involve gains and/or losses of TF binding sites (as opposed to modifications in the TF themselves (3–5), although modifications in TFs have also been reported (6–8)). The prevalence of differences in *cis*-regulatory sites has been ascribed to the fact that changes in the binding site for a particular TF can alter transcription of a single gene selectively without affecting the expression of other genes co-regulated by the same TF (3–5). Consistent with this notion, mutations in *cis*-regulatory elements underlie several morphological differences among closely related animal species (3, 5, 9, 10).

We hypothesized that the evolution of gene regulation in bacteria may differ from what has been described thus far in eukaryotes for several reasons. First, it has been suggested that TFs are more likely to evolve if they acquire new target genes (7) and, unlike closely related eukaryotic organisms, closely related bacterial species often exhibit substantial differences in gene content, so that orthologous TFs control largely distinct gene sets (11–15). Second, the position and orientation of a TF binding site within a bacterial promoter (i.e., the promoter architecture) are critical for gene transcription (16). By contrast, eukoryotic promoters are characterized by the sparse and uneven distribution of TF binding sites over large DNA regions (17).

To test our hypothesis, we examined the regulatory protein PhoP and its regulated targets in 2 related enteric bacteria: the gastroenteritis-causing *Salmonella enterica* serovar Typhimurium and the bubonic plague agent *Yersinia pestis*. The PhoP protein governs virulence and the adaptation to low-Mg^{2+} environments in these 2 bacterial pathogens (18, 19), and its activity is dictated by the Mg^{2+}-responding sensor PhoQ (20). Although the *Salmonella* and *Yersinia* PhoP proteins are 79% identical at the amino acid level, most of the genes directly regulated by PhoP in *S. enterica* have no homologs in *Y. pestis*, and vice versa (Perez, et al., unpublished results), probably as a result of horizontal gene transfer.

Here, we show that (*i*) the *Salmonella* and *Yersinia* PhoP proteins differ in their ability to promote transcription of certain species-specific genes although they retain the capacity to transcribe ancestral members of the PhoP regulon, and (*ii*) the *Salmonella* and *Yersinia* genomes each harbor promoter architectures not found in the other genome, but the PhoP protein recognizes a conserved motif in these 2 organisms. Our results suggest that embedding new target genes under the control of an ancestral TF entails changes in TF function and in promoter architecture.

Results

Differential Ability of the *Yersinia* PhoP Protein to Express *Salmonella* Genes Absent from *Yersinia*. To examine whether the *Yersinia* PhoP protein was functionally equivalent to the *Salmonella* PhoP protein, we used the experimental strategy schematized in supporting information (SI) Fig. S1. We began by investigating *Yersinia*'s ability to promote transcription of the *S. enterica mgtA* and *ugtL* genes, which have no orthologs in *Y. pestis* and harbor distinct promoter architectures (Fig. 1*A*) (21–23). *Y. pestis* produced fluorescence levels similar to those of *S. enterica* when transcription of a promoterless *gfp* gene was driven by the promoter of the *mgtA* gene (Fig. 1*B*), which is found in most enteric species (24). By contrast, *Yersinia* elicited significantly lower fluorescence than *Salmonella* when *gfp* was expressed from the promoter of the *Salmonella*-specific *ugtL* gene (Fig. 1*C*). In principle, *Yersinia*'s inability to promote *ugtL* transcription could be caused by the absence of a co-regulator(s) normally present in *Salmonella*. To test this possibility, we examined *ugtL* transcription in an engineered *Salmonella* strain in which the *phoPQ* operon was replaced by the *Y. pestis phoPQ* operon (Fig. 1*D*). The resulting strain synthesized the *Yersinia* PhoP protein under the same conditions and at similar levels as the *Salmonella* PhoP protein in the strain expressing *Salmonella*'s *phoPQ* operon (Fig. 1*E*). There was no PhoP-promoted *ugtL* transcription in the strain expressing the *Yersinia phoPQ* operon (Fig. 1*G*), even though the *mgtA* gene was tran-

Author contributions: J.C.P. and E.A.G. designed research; J.C.P. performed research; J.C.P. and E.A.G. analyzed data; and J.C.P. and E.A.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]Present address: Department of Microbiology and Immunology, University of California, 600 16th St., Genentech Hall Room N374, San Francisco, CA 94158.

[2]To whom correspondence should be addressed. E-mail: groisman@borcim.wustl.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0810343106/DCSupplemental.

GENETICS

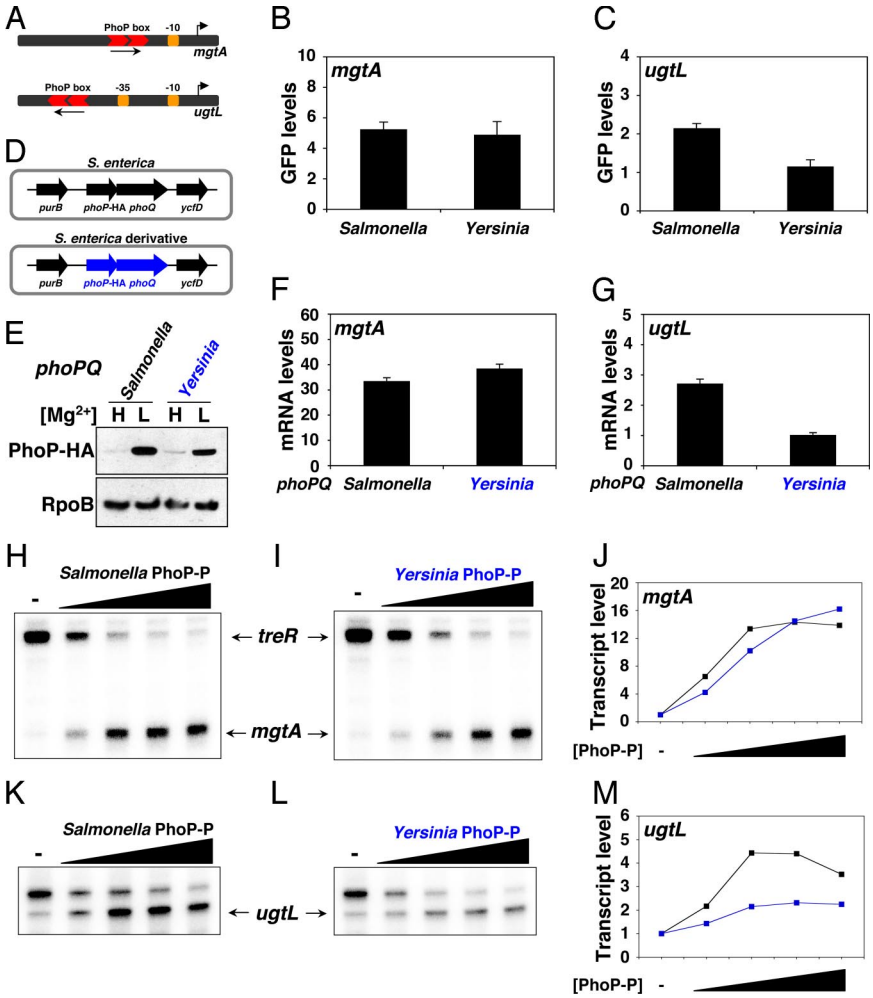


Fig. 1. The *Salmonella* and *Yersinia* PhoP proteins differ in their ability to promote gene transcription. (*A*) The *Salmonella mgtA* and *ugtL* promoters differ both in the orientation and the distance of the PhoP box to the promoter −10 region. The PhoP box is located in the place that corresponds to the promoter −35 region in the *mgtA* promoter but is 20 bp further upstream in the *ugtL* promoter. (*B* and *C*) GFP expression driven by the *Salmonella mgtA* and *ugtL* promoters in *Salmonella* and *Yersinia*. Approximately 150-nt DNA fragments corresponding to the *mgtA* and *ugtL* promoter regions (covering ≈130 nucleotides upstream and ≈20 nucleotides downstream of the transcription start sites) were cloned in front of a promoterless *gfp* gene in the low-copy plasmid pMS201. Organisms were grown in defined medium containing 50 μM $MgSO_4$, inducing conditions for the PhoP/PhoQ system. GFP expression was normalized to cell density. Shown are the ratios of the normalized GFP values between wild-type *Salmonella* (14028s) or *Yersinia* (KIM6) and their respective isogenic *phoP* mutant strains (EG15598 and EG14737, respectively). Values shown are mean plus SD of at least 3 independent experiments. (*D*) Schematic of the genomic context of the *phoPQ* locus in *Salmonella* strain EG13918 coding for the *Salmonella* PhoP-HA and PhoQ proteins and its derivative harboring the *Yersinia phoP*-HA and *phoQ* genes (in blue) (EG17569). (*E*) Expression of the *Salmonella* and *Yersinia* PhoP proteins in the *Salmonella* strains EG13918 and EG17569 depicted in *D*. Western blot analysis was performed with anti-HA and anti-RpoB antibodies (to detect the PhoP-HA and RpoB proteins, respectively) on cell extracts prepared from bacteria grown as described in *B* and *C* in medium containing 10 mM (H) or 50 μM (L) $MgSO_4$. (*F* and *G*) *mgtA* and *ugtL* expression in the *Salmonella* strains depicted in *D*. Cells were grown as described in *B* and *C*. Transcript levels were determined by quantitative real-time PCR and normalized to ribosomal RNA levels. Shown are the ratios of the normalized transcript levels present in the strains described in *D* relative to those produced by the *phoPQ* mutant EG15598. Values shown are mean plus SD of at least 3 independent experiments. (*H–M*) Single-round in vitro transcription assays with linear templates corresponding to the *mgtA* (*H–J*) and *ugtL* (*K–M*) promoters, *E. coli* RNA polymerase, and increasing amounts of phosphorylated *Salmonella* or *Yersinia* PhoP proteins. The upper band in *H–I* corresponds to *treR*, a PhoP-repressed transcript going in the reverse orientation (22). The upper band in *K* and *L* corresponds to a spurious transcript observed in vitro but not in vivo (21, 45). Quantification of the in vitro transcription assays is shown in *J* for *mgtA* and in *M* for *ugtL*.

scribed at similar levels as in the *Salmonella* strain harboring its own *phoPQ* operon (Fig. 1*F*).

We reasoned that if the observed differences in *ugtL* expression were caused by intrinsic properties of the PhoP proteins, it should be possible to recapitulate the observed behavior (Fig. 1 *B*, *C*, *F*, and *G*) in an in vitro transcription system where the only variable was the PhoP protein. Consistent with this notion, *ugtL* transcription in vitro was significantly higher with the *Salmonella* PhoP protein than with the *Yersinia* PhoP protein (Fig. 1 *K–M*). By contrast, the *Salmonella* and *Yersinia* PhoP proteins promoted similar levels of *mgtA* transcription in vitro (Fig. 1 *H–J*). Cumulatively, these results indicate that the function of the PhoP protein has been conserved only partially since *Salmonella* and *Yersinia* diverged from their last common ancestor > 200 million years ago (25).

The *Yersinia* PhoP Protein Binds to the *Salmonella ugtL* Promoter but Does Not Recruit RNA Polymerase. The inability of the *Yersinia* PhoP protein to transcribe the *ugtL* gene could be caused by the absence of binding to the *ugtL* promoter and/or could result from a defect in a subsequent step in transcription such as recruitment of RNA polymerase. We ruled out the first possibility for 2 reasons. First, the *Yersinia* PhoP protein bound to both the *ugtL* and *mgtA* promoter regions in vivo just like the *Salmonella* PhoP protein (Fig. 2 *A* and *B*). Second, the purified PhoP proteins from *Yersinia* and

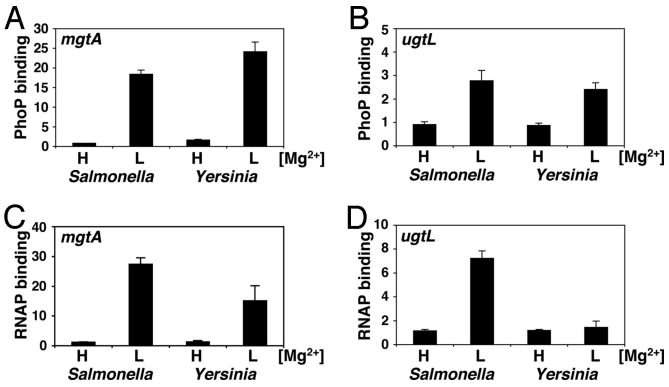


Fig. 2. The *Yersinia* PhoP protein binds to the *ugtL* promoter but cannot recruit RNA polymerase. (*A–D*) Promoter occupancy by PhoP (*A* and *B*) and RNA polymerase (*C* and *D*) determined by ChIP in isogenic *Salmonella* strains expressing the *Salmonella* (EG13918) or the *Yersinia* (EG17569) *phoPQ* operon. Cells were grown in *N*-minimal medium containing 10 mM (H) or 10 μM (L) $MgCl_2$. Shown are the mean plus SD of at least 3 independent experiments.

Salmonella displayed similar affinities toward a fragment harboring the *ugtL* promoter in vitro (Fig. S2). As expected, the PhoP protein occupied the *ugtL* and *mgtA* promoters only when grown under low Mg^{2+}-inducing conditions for the PhoP/PhoQ system (Fig. 2 *A* and *B*).

We then tested the ability of the *Salmonella* and *Yersinia* PhoP proteins to recruit RNA polymerase to the *ugtL* and *mgtA* promoters in vivo using ChIP with an antibody directed to the β subunit of the enzyme. RNA polymerase bound to both the *mgtA* and *ugtL* promoter regions when the *Salmonella* strain expressing its own *phoP* gene was grown in the presence of the low Mg^{2+}-inducing signal (Fig. 2 *C* and *D*). By contrast, RNA polymerase could be recruited to the *mgtA* promoter but not to the *ugtL* promoter in the *Salmonella* strain expressing the *Yersinia phoP* gene (Fig. 2 *C* and *D*). As expected, there was no RNA polymerase recruitment to either promoter in either strain under high Mg^{2+}-repressing conditions (Fig. 2 *C* and *D*). Taken together, these findings indicate that the *Yersinia* PhoP protein fails to promote *ugtL* transcription because of inefficient RNA polymerase recruitment rather than because of defective binding to the *ugtL* promoter.

Promoter Architecture Dictates the Ability of the *Salmonella* and *Yersinia* PhoP Proteins to Transcribe Target Genes. The orientation and/or location of the PhoP box seem to be responsible (at least in part) for the distinct abilities of the *Salmonella* and *Yersinia* PhoP proteins to transcribe the *mgtA* and *ugtL* genes, because the corresponding promoters differ in both of these properties (Fig. 1*A*). In agreement with this notion, the *Salmonella* strain expressing the *Yersinia phoPQ* operon transcribed the ancestral *slyB* (Fig. S3) and *phoP* (Fig. 1*E*) genes, whose promoters share the location and orientation of the PhoP box with the *mgtA* promoter (Fig. 3*A*) but not the horizontally acquired *pagK*, *mgtC*, *pagC*, and *mig-14* genes (Fig. S3 and data not shown), the promoters of which contain the PhoP box in the same orientation as in the *ugtL* promoter (Fig. 3*A*). This finding suggested to us that the architecture of natural PhoP-activated promoters in *Yersinia* would resemble the *Salmonella mgtA* promoter architecture and not conform to the *ugtL* promoter structure.

To determine the architecture of PhoP-activated promoters in *Yersinia*, we identified the transcription start site (by conducting S1 mapping or primer extension experiments) and defined the PhoP binding site (by DNase I footprinting experiments with the *Yersinia* PhoP protein) for 14 *Y. pestis* genes directly controlled by the PhoP protein (Figs. 3*B* and S4) (see *Materials and Methods* for a description of how these genes were uncovered). Analysis of the sequence, location, and orientation of the PhoP box revealed interesting similarities and differences between PhoP-activated promoters in *Salmonella* and *Yersinia*. First, a comparison with previously reported *Salmonella* PhoP-regulated promoters (Fig. 3*A*) (23, 26) revealed that the sequence of the PhoP-binding motif is similar in *Salmonella* and *Yersinia* (Fig. S5). Second, 9 of 14 *Yersinia* promoters contain a PhoP box in the same orientation and at the same distance from the −10 region as in the *Salmonella mgtA* promoter. And third, the remaining 5 *Yersinia* promoters also harbor the PhoP box in the same orientation as in the *Salmonella mgtA* promoter but further upstream: for 4 of these promoters the PhoP box is located at the same distance from the −10 region as in a group of *Salmonella* PhoP-activated promoters (i.e., *rstA*, *pagP*, and *ompX*). Therefore, none of the *Yersinia* PhoP-activated promoters has the PhoP box in the orientation found in the *Salmonella ugtL* promoter.

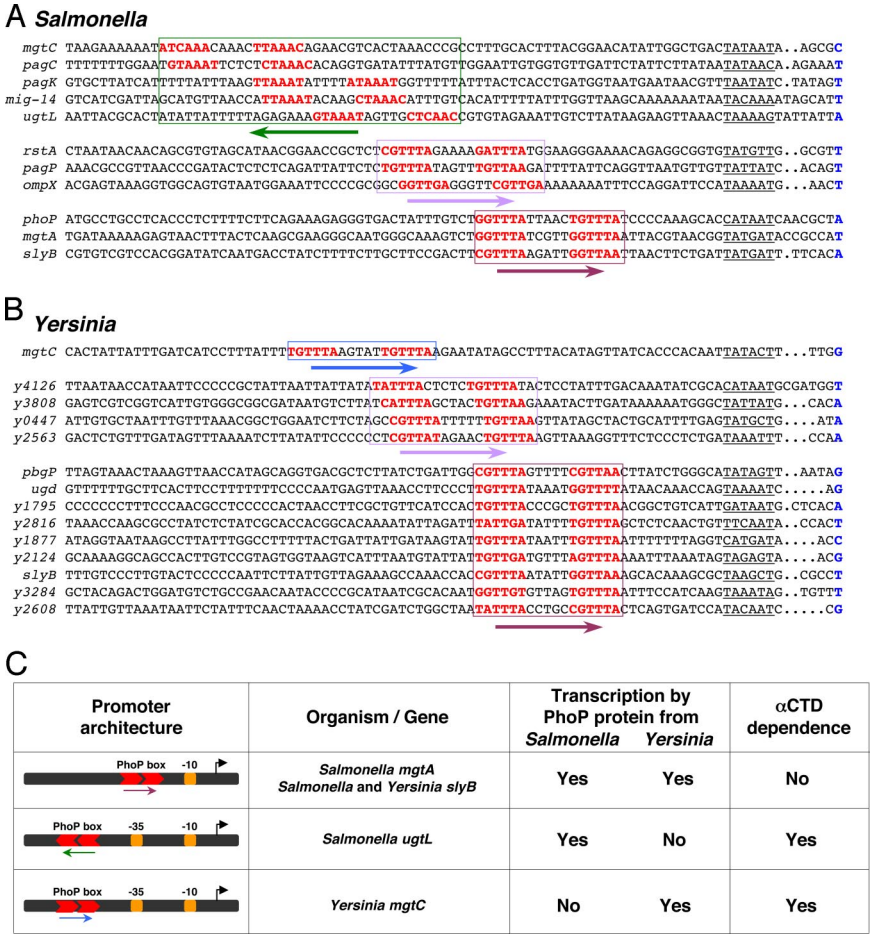
The promoter of the *Yersinia mgtC* gene exhibits an architecture not previously reported for *Salmonella* PhoP-activated promoters: the PhoP box is at a similar distance from the −10 region as the *Salmonella ugtL* promoter but is in the opposite orientation (Fig. 3 *B* and *C*). Thus, we wondered whether the *Salmonella* PhoP protein could promote transcription from the *Yersinia mgtC* promoter (the *Salmonella* and *Yersinia mgtC* genes are xenologs rather than orthologs, because they seem to have been acquired independently by the *Salmonella* and *Yersinia* lineages) (24, 27). The *Salmonella* PhoP protein could not promote transcription of the *Yersinia mgtC* gene (Fig. S6), but it elicited normal transcription levels of the *Yersinia y1795* gene (data not shown), which has the *Salmonella mgtA*-like promoter architecture (Fig. 3 *B* and *C*). These results suggest that the ability of the PhoP protein to promote gene transcription is dependent on the orientation of the PhoP box.

The αCTD Subunit of RNA Polymerase Is Necessary for PhoP-Dependent Transcription of *ugtL* but Not *mgtA*. The location of a TF binding site within a bacterial promoter determines the RNA polymerase subunit (e.g., σ or αCTD for the carboxy-terminal domain of the α subunit) with which an activator interacts to promote gene transcription (16). In class I promoters, the activator binds upstream of the −35 element in a promoter, and activation typically relies on interactions with the flexible αCTD. In class II promoters, the activator binds to a target that overlaps the promoter −35 element usually contacting the σ subunit of RNA polymerase (16). Because the αCTD is essential for viability (28), we used in vitro transcription assays with RNA polymerase reconstituted with either wild-type or truncated αCTD (29) to probe the requirement of αCTD for expression of PhoP-regulated promoters. The αCTD was dispensable for PhoP-dependent *mgtA* transcription (Fig. S7), as expected for a class II promoter and consistent with previous findings reported for the *Escherichia coli mgtA* promoter and the *E. coli* PhoP protein (22). By contrast, the αCTD was necessary for PhoP-dependent transcription of the class I *ugtL* promoter (Fig. S7).

A *Salmonella* Strain Expressing the *Yersinia phoPQ* Operon in Lieu of Its Own *phoPQ* Operon Is Attenuated for Virulence. A functional PhoP/PhoQ system is required for virulence in both *Salmonella* and *Yersinia* (18, 19). However, a *Salmonella* strain expressing the *Yersinia phoPQ* operon was strongly attenuated for virulence, because all the mice inoculated with a dose > 1,000 higher than the LD_{50} survived, as did those inoculated with a *phoP* null mutant, whereas those inoculated with a *Salmonella* strain harboring its own *phoPQ* operon died (data not shown).

Discussion

It is becoming increasingly clear that transcriptional regulatory circuits undergo major modifications over time and that these modifications provide an important source of phenotypic diversity. Although several studies have addressed the evolution of gene



C

Promoter architecture	Organism / Gene	Transcription by PhoP protein from *Salmonella*	Transcription by PhoP protein from *Yersinia*	αCTD dependence
PhoP box -10	*Salmonella mgtA* *Salmonella* and *Yersinia slyB*	Yes	Yes	No
PhoP box -35 -10	*Salmonella ugtL*	Yes	No	Yes
PhoP box -35 -10	*Yersinia mgtC*	No	Yes	Yes

Fig. 3. Architectures of PhoP-activated promoters in *Salmonella* and *Yersinia*. (*A* and *B*) Sequences of the promoter regions of the transcripts directly activated by PhoP in *Salmonella* (*A*) and *Yersinia* (*B*). The transcription start sites are indicated in blue, the PhoP boxes are in red, and the putative −10 sequences are underlined. Boxes and arrows indicate the location and orientation of the PhoP binding sites across promoters. The *Salmonella* PhoP-activated transcripts have been described previously (23, 26). The *Yersinia* PhoP-dependent transcripts were uncovered as described in *Materials and Methods* (Fig. S4). PhoP binding to all promoters was determined in vivo by ChIP-chip and/or ChIP-real-time PCR and in several cases in vitro by DNase I footprinting (Fig. S4). (*C*) Summary of the distribution of promoter architectures and the differential ability of the *Salmonella* and *Yersinia* PhoP proteins to activate transcription. Although the *mgtA*-like promoter architecture is shared by several *Salmonella* and *Yersinia* PhoP-activated genes, other members of the regulon harbor promoter structures that are absent from the other organism. The *Salmonella* and *Yersinia* PhoP proteins differ in their ability to activate expression from genes of the latter group.

regulatory circuits in eukaryotes, little is known about this process in bacteria. Here, we report that the evolution of the regulons controlled by the PhoP protein in the enteric pathogens *S. enterica* and *Y. pestis* has entailed functional changes in the PhoP protein as well as in the architecture of PhoP-dependent promoters, with conservation of the DNA motif recognized by the PhoP protein. Several lines of evidence support this notion. First, the *Salmonella* and *Yersinia* PhoP proteins differ in their ability to promote transcription of subsets of PhoP-regulated genes from the other species (Figs. 1, S3, and S6). Second, *Salmonella* and *Yersinia* harbor PhoP-regulated promoter architectures not found in the other species (Fig. 3). Third, the sequence of the PhoP binding motif is similar in *Salmonella* and *Yersinia* (Fig. S5). Fourth, a *Salmonella* strain expressing the *Yersinia phoPQ* operon in lieu of its own *phoPQ* operon is attenuated for virulence (data not shown). Our data provide a singular example of functional differences between orthologous TFs and between the architectures of their target promoters across bacterial species.

The PhoP proteins from *Salmonella* and *Yersinia* are fully competent to express genes harboring the conserved promoter architecture driving transcription of the ancestral members of the PhoP regulon but differ in their ability to transcribe species-specific genes harboring promoter architectures not found in the other species (Figs. 1 and 3). PhoP-regulated promoters can be divided into 3 groups based on the location of the PhoP box (Fig. 3 *A* and *B*): (*i*) the prototypical class II promoter (16), exemplified by *mgtA* and the ancestral members of the regulon *phoP* and *slyB* (Perez, et al., unpublished results), in which the PhoP box is ≈12 nucleotides from the −10 region; (*ii*) promoters in which the PhoP box is located further upstream but in the same orientation as in the *mgtA* promoter (e.g., the *Salmonella rstA*, *pagP*, and *ompX* promoters and the *Yersinia y4126*, *y3808*, *y0447*, and *y2563* promoters); and (*iii*), the typical class I promoters that contain the PhoP box even further upstream but at a variable distance from the −10 region (e.g., the *Salmonella ugtL* and *mgtC* promoters and the *Yersinia mgtC* promoter). The first 2 promoter architectures are present both in *Salmonella* and *Yersinia*, and the PhoP proteins from these 2 species can transcribe these promoters regardless of the species of origin (Figs. 1 *B*, *F*, *H*, and *I*, S3, and data not shown). The orientation of the PhoP box in the third group of promoters is different in *Salmonella* and in *Yersinia* (Fig. 3*C*). Strikingly, this difference seems to be critical for PhoP function, because the *Yersinia* PhoP protein is impaired in its ability to elicit transcription from the

Salmonella promoters of this group, and vice versa (Figs. 1 *C*, *G*, and *L*, 3*C*, and S6). The *Salmonella* genes harboring the latter promoters have no close homologs outside *Salmonella* spp., suggesting that they were horizontally acquired by the lineage that gave rise to *S. enterica*.

The surface of a TF that contacts RNA polymerase is determined by the orientation and location of a TF binding site (30). For instance, the transcriptional activator cyclic APM receptor protein from *E. coli* interacts with αCTD through more than 1 surface, depending on the position of its binding site (31). The fact that αCTD is dispensable for transcription of PhoP-activated class II promoters (e.g., *mgtA*) but is necessary to transcribe the 2 groups of class I promoters (e.g., *ugtL*, *rstA*, and *Yersinia mgtC*) (Fig. S7 and data not shown) suggests that the *Salmonella* and *Yersinia* PhoP proteins interact with both σ and with αCTD, depending on the location of the PhoP box. The contact(s) that PhoP establishes with σ in the first group of promoters (e.g., *mgtA*) seems to be conserved between the 2 PhoP proteins, because *Salmonella* and *Yersinia* can promote transcription from this type of promoter. Similarly, the potential interaction(s) of PhoP with αCTD (and perhaps with σ as well) when transcribing from the second group of promoters (e.g., the *Salmonella rstA*, *pagP*, and *ompX* promoters and the *Yersinia y4126*, *y3808*, *y0447*, and *y2563* promoters) is likely to be conserved between *Salmonella* and *Yersinia* because these promoters share the position and orientation of the PhoP box relative to the −10 region. By contrast, the *Salmonella* and *Yersinia* PhoP proteins seem to establish productive interactions with αCTD in different manners when transcribing the third group of promoters (e.g., the *Salmonella ugtL* and *mgtC* promoter and the *Yersinia mgtC* promoter) that differ in the orientation of the PhoP box. Therefore, although the known flexibility of αCTD (31) enables the PhoP protein to transcribe promoters harboring the PhoP box at a variety of positions (green box in Fig. 3*A*), certain promoter architectures seem to be species specific and function only with the PhoP protein corresponding to that species. Because the αCTD subunit of *E. coli*, *S. enterica*, and *Y. pestis* are 100% identical at the amino acid level, the functional differences between the *Salmonella* and *Yersinia* PhoP proteins cannot be ascribed to differences in the transcription machinery.

The evolution of certain transcriptional regulators in yeast has been ascribed to changes in protein–protein interactions caused by a few amino acid substitutions (1, 2, 32), whereas the reported cases in higher eukaryotes typically involve whole-domain gains or losses (6). (However, some examples in which a few amino acid substitutions seem to be responsible for functional changes have also been reported (7, 33).) The *Salmonella* and *Yersinia* PhoP proteins have 2 domains: a 120-residue *N*-terminal domain joined by a 5-residue linker to a 98-residue *C*-terminal domain (34). (The *N*- and *C*-terminal domains are 80.8% and 75.5% identical, respectively.) The *C*-terminal region is predicted to make contacts with RNA polymerase through unidentified residues and has a helix-turn-helix DNA binding domain. The helix-turn-helix portion is identical in *Salmonella* and *Yersinia* proteins, in agreement with the conservation of the DNA motif that PhoP recognizes (Fig. S5). Moreover, the predicted secondary structures of the entire *Salmonella* and *Yersinia* PhoP proteins are superimposable (Fig. S8), suggesting that the functional differences between these orthologous PhoP proteins result from relatively minor changes in the overall structure of the proteins altering the interaction with the transcriptional machinery.

A *Salmonella* strain expressing the *Yersinia phoPQ* operon was strongly attenuated for virulence, behaving similarly to a *phoP* null mutant (data not shown). Given the degree of sequence identity between the sensing domains of the *Salmonella* and *Yersinia* PhoQ proteins (35), the virulence attenuation of the *Salmonella* strain expressing the *Yersinia phoPQ* operon is unlikely to result from a defect in the sensing ability of the *Yersinia* PhoQ protein. Rather, it may reflect that *Salmonella* virulence demands the expression of horizontally acquired PhoP-dependent genes, such as *mig-14* (36) and *mgtC* (24), whose promoters share the location and orientation of the PhoP box with the *ugtL* promoter (Fig. 3*A*) and thus are not efficiently transcribed by the *Y. pestis* PhoP protein (data not shown).

In sum, the evolution of a bacterial regulon entails modifications in a TF and restructuring of its target promoters, which enable transcription of horizontally acquired targets while retaining control of ancestral genes. These changes are in addition to the well-established transcriptional rewiring events whereby the promoters of conserved genes in related species differ in the presence/absence of TF binding site(s) (1, 37, 38). While there are a few reports of evolutionary changes in eukaryotic TFs contributing to the modification of gene-regulatory networks (6–8, 33), alterations in gene regulation and phenotypic differences in eukaryotes have been ascribed primarily to gains and/or losses of TF binding sites (3, 5, 10). Our results (Fig. 2) and those of others (39) indicate that, in bacteria, TF binding to a DNA sequence in vivo does not necessarily result in gene transcription, thereby highlighting the need to experimentally verify the function of identified TF binding sites.

Materials and Methods

Bacterial Strains, Plasmids, and Growth Conditions. Bacterial strains and plasmids used in this study are listed in Table S1 and primers are listed in Table S2. All *Salmonella enterica* serovar Typhimurium strains were derived from wild-type strain 14028s and were grown at 37 °C in N-minimal medium (40) buffered in 50 mM Bis-Tris, pH 7.7, supplemented with 0.1% casamino acids, 38 mM glycerol, and 10 μM or 10 mM $MgCl_2$, unless otherwise indicated. *Yersinia pestis* strains were derived from wild-type strain KIM6 (41) and were grown at the optimal growth temperature of 28 °C in a modified defined medium (42), pH 7.0, supplemented with 0.1% casamino acids, 10 mM (D)-glucosamine, and 50 μM or 10 mM $MgSO_4$. *Escherichia coli* strain DH5α was used as the host for the preparation of plasmid DNA. Ampicillin and kanamycin were used at 50 μg/ml and chloramphenicol at 20 μg/ml. A detailed description of the construction of bacterial strains and plasmids as well as other molecular biology procedures is provided in the *SI Text*.

Identification of PhoP-Regulated Promoters in *Y. pestis*. To uncover the genes directly regulated by PhoP in *Y. pestis*, we used a combination of expression microarray analysis (wild-type vs. *phoP* mutant strains) and ChIP followed by array hybridization (ChIP-chip). S1 mapping (43), primer extension (44), and DNase I footprinting (43) experiments were used to validate the microarray and ChIP-chip results and to determine the structure of the PhoP-activated promoters described in Fig. 3*B*. A comprehensive description of the procedures and the uncovered genes will be published elsewhere.

GFP Expression Analysis. After overnight culture in defined medium containing 10 mM $MgSO_4$, the *Salmonella* and *Yersinia* strains were washed twice with Mg^{2+}-free medium, inoculated (1:50 dilution) in medium containing 50 μM $MgSO_4$, and grown to mid-exponential phase with vigorous shaking. Fluorescence values were normalized to cell density, both of which were measured in a Victor^3 1420 Multilabel counter (Perkin Elmer).

RNA Isolation and Real-Time PCR to Determine Transcript Levels. *Salmonella* cells were grown as described in the previous paragraphs for the GFP expression analysis. Total RNA was prepared using the SV Total RNA Isolation System (Promega). cDNA was synthesized using TaqMan (Applied Biosystems) and random hexamers following the manufacturer's instructions. Quantification of transcripts was performed by real-time PCR using SYBR Green PCR Master Mix (Applied Biosystems) in an ABI 7000 Sequence Detection System (Applied Biosystems). Results were normalized to the levels of 16S ribosomal RNA.

ACKNOWLEDGMENTS. We thank Richard L. Gourse for providing the mutant RNA polymerase enzyme; S. Busby, B. Cohen, R.L. Gourse, and G. Wagner for critically reading the manuscript; and Tammy Latifi for technical assistance. The work presented here was supported, in part, by Grant AI49561 from the National Institutes of Health to E.A.G, who is an Investigator of the Howard Hughes Medical Institute.

1. Tuch BB, Galgoczy DJ, Hernday AD, Li H, Johnson AD (2008) The evolution of combinatorial gene regulation in fungi. *PLoS Biol* 6:e38.
2. Tuch BB, Li H, Johnson AD (2008) Evolution of eukaryotic transcription circuits. *Science* 319:1797–1799.
3. Carroll SB (2005) Evolution at two levels: On genes and form. *PLoS Biol* 3:e245.
4. Stern DL (2000) Evolutionary developmental biology and the problem of variation. *Evolutio (Lawrence, Kans)* 54:1079–1091.
5. Wray GA (2007) The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet* 8:206–216.
6. Hsia CC, McGinnis W (2003) Evolution of transcription factor function. *Curr Opin Genet Dev* 13:199–206.
7. Wagner GP, Lynch VJ (2008) The gene regulatory logic of transcription factor evolution. *Trends Ecol Evol* 23:377–385.
8. Lynch VJ, Wagner GP (2008) Resurrecting the role of transcription factor change in developmental evolution. *Evolution* 62:2131–2154.
9. Jeong S, et al. (2008) The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* 132:783–793.
10. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* 134:25–36.
11. Lozada-Chavez I, Janga SC, Collado-Vides J (2006) Bacterial regulatory networks are extremely flexible in evolution. *Nucleic Acids Res* 34:3434–3445.
12. Madan Babu M, Teichmann SA, Aravind L (2006) Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J Mol Biol* 358:614–633.
13. Price MN, Dehal PS, Arkin AP (2007) Orthologous transcription factors in bacteria have different functions and regulate different genes. *PLOS Comput Biol* 3:1739–1750.
14. Rhodius VA, Suh WC, Nonaka G, West J, Gross CA (2006) Conserved and variable functions of the σ^E stress response in related genomes. *PLoS Biol* 4:e2.
15. Swingle B, et al. (2008) Characterization of the PvdS-regulated promoter motif in *Pseudomonas syringae* pv. tomato DC3000 reveals regulon members and insights regarding PvdS function in other pseudomonads. *Mol Microbiol* 68:871–889.
16. Browning DF, Busby SJ (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2:57–65.
17. Wray GA, et al. (2003) The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* 20:1377–1419.
18. Groisman EA (2001) The pleiotropic two-component regulatory system PhoP-PhoQ. *J Bacteriol* 183:1835–1842.
19. Oyston PC, et al. (2000) The response regulator PhoP is important for survival under conditions of macrophage-induced stress and virulence in *Yersinia pestis. Infect Immun* 68:3419–3425.
20. Garcia Vescovi E, Soncini FC, Groisman EA (1996) Mg^{2+} as an extracellular signal: Environmental regulation of *Salmonella* virulence. *Cell* 84:165–174.
21. Shi YX, Latifi T, Cromie MJ, Groisman EA (2004) Transcriptional control of the antimicrobial peptide resistance *ugtL* gene by the *Salmonella* PhoP and SlyA regulatory proteins. *J Biol Chem* 279:38618–38625.
22. Yamamoto K, Ogasawara H, Fujita N, Utsumi R, Ishihama A (2002) Novel mode of transcription regulation of divergently overlapping promoters by PhoP, the regulator of two-component system sensing external magnesium availability. *Mol Microbiol* 45:423–238.
23. Zwir I, et al. (2005) Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica. Proc Natl Acad Sci USA* 102:2862–2867.
24. Blanc-Potard AB, Groisman EA (1997) The *Salmonella selC* locus contains a pathogenicity island mediating intramacrophage survival. *EMBO J* 16:5376–5385.
25. Deng W, et al. (2002) Genome sequence of *Yersinia pestis* KIM. *J Bacteriol* 184:4601–4611.
26. Lejona S, Aguirre A, Cabeza ML, Vescovi EG, Soncini FC (2003) Molecular characterization of the Mg^{2+}-responsive PhoP-PhoQ regulon in *Salmonella enterica. J Bacteriol* 185:6287–6294.
27. Blanc-Potard AB, Lafay B (2003) MgtC as a horizontally-acquired virulence factor of intracellular bacterial pathogens: Evidence from molecular phylogeny and comparative genomics. *J Mol Evol* 57:479–486.
28. Gaal T, et al. (1996) DNA-binding determinants of the α subunit of RNA polymerase: Novel DNA-binding domain architecture. *Genes Dev* 10:16–26.
29. Quinones M, Kimsey HH, Ross W, Gourse RL, Waldor MK (2006) LexA represses CTXΦ transcription by blocking access of the α C-terminal domain of RNA polymerase to promoter DNA. *J Biol Chem* 281:39407–39412.
30. Martin RG, Gillette WK, Rhee S, Rosner JL (1999) Structural requirements for marbox function in transcriptional activation of mar/sox/rob regulon promoters in *Escherichia coli*: Sequence, orientation and spatial relationship to the core promoter. *Mol Microbiol* 34:431–441.
31. Niu W, Kim Y, Tau G, Heyduk T, Ebright RH (1996) Transcription activation at class II CAP-dependent promoters: Two interactions between CAP and RNA polymerase. *Cell* 87:1123–1134.
32. Tsong AE, Tuch BB, Li H, Johnson AD (2006) Evolution of alternative transcriptional circuits with identical logic. *Nature* 443:415–420.
33. Lynch VJ, et al. (2008) Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. *Proc Natl Acad Sci USA* 105:14928–14933.
34. Bachhawat P, Stock AM (2007) Crystal structures of the receiver domain of the response regulator PhoP from *Escherichia coli* in the absence and presence of the phosphoryl analog beryllofluoride. *J Bacteriol* 189:5987–5995.
35. Prost LR, Daley ME, Bader MW, Klevit RE, Miller SI (2008) The PhoQ histidine kinases of *Salmonella* and *Pseudomonas* spp. are structurally and functionally different: Evidence that pH and antimicrobial peptide sensing contribute to mammalian pathogenesis. *Mol Microbiol* 69:503–519.
36. Brodsky IE, Ghori N, Falkow S, Monack D (2005) Mig-14 is an inner membrane-associated protein that promotes *Salmonella typhimurium* resistance to CRAMP, survival within activated macrophages and persistent infection. *Mol Microbiol* 55:954–972.
37. Odom DT, et al. (2007) Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39:730–732.
38. Borneman AR, et al. (2007) Divergence of transcription factor binding sites across related yeast species. *Science* 317:815–819.
39. Shimada T, Ishihama A, Busby SJ, Grainger DC (2008) The *Escherichia coli* RutR transcription factor binds at targets within genes as well as intergenic regions. *Nucleic Acids Res* 36:3950–3955.
40. Snavely MD, Gravina SA, Cheung TT, Miller CG, Maguire ME (1991) Magnesium transport in *Salmonella Typhimurium*—regulation of Mgta and Mgtb expression. *J Biol Chem* 266:824–829.
41. Staggs TM, Perry RD (1991) Identification and cloning of a fur regulatory gene in *Yersinia pestis. J Bacteriol* 173:417–425.
42. Perry RD, Brubaker RR (1979) Accumulation of iron by *Yersiniae*. *J Bacteriol* 137:1290–1298.
43. Winfield MD, Latifi T, Groisman EA (2005) Transcriptional regulation of the 4-amino-4-deoxy-L-arabinose biosynthetic genes in *Yersinia pestis. J Biol Chem* 280:14765–14772.
44. Tu X, Latifi T, Bougdour A, Gottesman S, Groisman EA (2006) The PhoP/PhoQ two-component system stabilizes the alternative sigma factor RpoS in *Salmonella enterica. Proc Natl Acad Sci USA* 103:13503–13508.
45. Perez JC, Latifi T, Groisman EA (2008) Overcoming H-NS-mediated transcriptional silencing of horizontally acquired genes by the PhoP and SlyA proteins in *Salmonella enterica. J Biol Chem* 283:10773–10783.