

# Mapping the Regulatory Network for *Salmonella enterica* Serovar Typhimurium Invasion

Carol Smith,<sup>a</sup> Anne M. Stringer,<sup>a</sup> Chunhong Mao,<sup>b</sup> Michael J. Palumbo,<sup>a</sup> Joseph T. Wade<sup>a,c</sup>

Wadsworth Center, New York State Department of Health, Albany, New York, USA<sup>a</sup>; Biocomplexity Institute of Virginia Tech, Virginia Tech, Blacksburg, Virginia, USA<sup>b</sup>; Department of Biomedical Sciences, School of Public Health, University at Albany, Albany, New York, USA<sup>c</sup>

C.S., A.M.S., and C.M. contributed equally to this work.

**ABSTRACT** *Salmonella enterica* pathogenicity island 1 (SPI-1) encodes proteins required for invasion of gut epithelial cells. The timing of invasion is tightly controlled by a complex regulatory network. The transcription factor (TF) HilD is the master regulator of this process and senses environmental signals associated with invasion. HilD activates transcription of genes within and outside SPI-1, including six other TFs. Thus, the transcriptional program associated with host cell invasion is controlled by at least 7 TFs. However, very few of the regulatory targets are known for these TFs, and the extent of the regulatory network is unclear. In this study, we used complementary genomic approaches to map the direct regulatory targets of all 7 TFs. Our data reveal a highly complex and interconnected network that includes many previously undescribed regulatory targets. Moreover, the network extends well beyond the 7 TFs, due to the inclusion of many additional TFs and noncoding RNAs. By comparing gene expression profiles of regulatory targets for the 7 TFs, we identified many uncharacterized genes that are likely to play direct roles in invasion. We also uncovered cross talk between SPI-1 regulation and other regulatory pathways, which, in turn, identified gene clusters that likely share related functions. Our data are freely available through an intuitive online browser and represent a valuable resource for the bacterial research community.

**IMPORTANCE** Invasion of epithelial cells is an early step during infection by *Salmonella enterica* and requires secretion of specific proteins into host cells via a type III secretion system (T3SS). Most T3SS-associated proteins required for invasion are encoded in a horizontally acquired genomic locus known as *Salmonella* pathogenicity island 1 (SPI-1). Multiple regulators respond to environmental signals to ensure appropriate timing of SPI-1 gene expression. In particular, there are seven transcription regulators that are known to be involved in coordinating expression of SPI-1 genes. We have used complementary genome-scale approaches to map the gene targets of these seven regulators. Our data reveal a highly complex and interconnected regulatory network that includes many previously undescribed target genes. Moreover, our data functionally implicate many uncharacterized genes in the invasion process and reveal cross talk between SPI-1 regulation and other regulatory pathways. All datasets are freely available through an intuitive online browser.

Received 9 June 2016 Accepted 10 August 2016 Published 6 September 2016

**Citation** Smith C, Stringer AM, Mao C, Palumbo MJ, Wade JT. 2016. Mapping the regulatory network for *Salmonella enterica* serovar Typhimurium invasion. *mBio* 7(5):e01024-16. doi:10.1128/mBio.01024-16.

**Editor** Stanley Maloy, San Diego State University

**Copyright** © 2016 Smith et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Joseph T. Wade, [joseph.wade@health.ny.gov](mailto:joseph.wade@health.ny.gov).

*Salmonella enterica* is the causative agent of typhoid fever and is also a major cause of foodborne illness (salmonellosis) (1). There are many serovars of *S. enterica* that cause salmonellosis; one of the clinically most important serovars is *S. enterica* subsp. *enterica* serovar Typhimurium (2). During the initial stages of infection, *S. Typhimurium* invades gut epithelial cells. The invasion process requires injection of specific effector proteins into host cells through a type III secretion system (T3SS). This T3SS and most of the secreted effector proteins are encoded in *Salmonella* pathogenicity island 1 (SPI-1), a horizontally acquired chromosomal region of ~40 kbp (3). Under standard laboratory growth conditions, SPI-1 genes are transcriptionally repressed. However, during the initial stages of infection, SPI-1 genes are induced in response to environmental triggers. These triggers can be mimicked in the laboratory by growth in media containing high levels of salt, by low levels of aeration, and by growth to the

late exponential/early stationary phase (3–6). The master regulator of SPI-1 genes is HilD, an AraC family transcription factor (TF) encoded within SPI-1 (7, 8). HilD expression and activity are controlled by multiple pathways that sense the environmental cues associated with invasion (8, 9). HilD activates transcription of several SPI-1 genes, including components of the T3SS, secreted effector proteins (10–12), and the TFs HilA and InvF (7, 10, 11). HilA and InvF activate transcription of additional T3SS components and effector proteins (8, 13–16). Approximately half of the known HilD-regulated genes are located outside SPI-1 (11). Notably, HilD activates transcription of genes within *Salmonella* pathogenicity island 4 (SPI-4), which are required for attachment of *S. Typhimurium* to epithelial cells (11, 17). HilD also activates transcription of genes whose function has not been directly connected to invasion, such as *lpxR*, which encodes an enzyme that modifies lipid A (11), and *flhDC*, which encodes the master regu-

lator of flagellar motility (11, 18, 19). Regulation of genes not involved in invasion may be a mechanism to coordinate their expression with the invasion process (11) or may indicate an as-yet-unidentified role in the invasion process.

In addition to HilD, HilA, and InvF, four other TFs within SPI-1, or connected to SPI-1, have been associated with host cell invasion. HilC and RtsA are close homologues of HilD, and their regulons overlap that of HilD (7, 8, 10, 11, 20, 21). HilC is encoded within SPI-1, whereas RtsA is encoded elsewhere in the genome. HilD, HilC, and RtsA not only regulate shared target genes but also activate each other's transcription (8, 10, 11, 20). Thus, these proteins represent a positive-feedback loop for activation of SPI-1 genes (20, 22). SprB is a TF encoded within SPI-1 also; however, little is known about the contribution of SprB to global gene regulation (23). Transcription of *sprB* is induced under conditions associated with invasion (4, 23), although the mechanism of this regulation is unknown. Only one regulatory target has been described for SprB: the operon that encompasses SPI-4 (24). RtsB is a TF encoded outside SPI-1 but connected to SPI-1 due to its regulation by HilD/HilC/RtsA (10). Moreover, RtsB is encoded in an operon with RtsA. The only known regulatory target of RtsB is *flhDC* (10, 19). However, unlike HilD, RtsB is a negative regulator of *flhDC* (10, 19). Thus, flagellar motility is regulated both positively and negatively by SPI-1-associated TFs (19). Although the precise functions of the 7 TFs described above have not been described, for simplicity we refer to them as "SPI-1-associated TFs."

Levels of active HilD, HilC, RtsA, and HilA are controlled by many different regulators, indirectly impacting expression of other SPI-1 genes (9). Nucleoid-associated proteins (NAPs) are abundant DNA-binding proteins that bind large numbers of genomic locations, typically with little sequence specificity (25). Several NAPs play crucial roles in regulating SPI-1 gene expression. For example, H-NS binds extensively within SPI-1 (26, 27), leading to repression of many SPI-1 genes, including *hilD*, *hilC*, *rtsA*, and *hilA* (28–31). Two other NAPs, Fis and HU, positively regulate SPI-1 genes, although the mechanism of regulation is unknown and may be indirect (32–35). Direct, positive regulation of SPI-1 genes is often due to displacement of H-NS by other DNA-binding proteins ("countersilencing" [36]). Indeed, this is the case for activation of *hilD*, *hilC*, *rtsA*, and *hilA* by HilD/HilC/RtsA (29, 30). Another NAP, IHF, also positively regulates *hilA* by displacing H-NS (31). In contrast, several transcription factors positively or negatively regulate SPI-1 genes by controlling *hilD* transcription independently of H-NS (e.g., Fur) (37, 38), or by controlling HilD translation (e.g., SirA and CsrA) (39) or activity (e.g., HilE and FliZ) (40, 41). Transcription factors also control SPI-1 gene expression by regulating *hilC* (OmpR) (35, 42) or *hilA* (PhoP and FNR) (5, 9, 43). By controlling expression/activity of HilD, HilC, RtsA, and HilA, many regulators contribute to SPI-1 gene regulation in response to environmental stimuli that include temperature (H-NS and Hha), osmolarity (H-NS and Hha), pH (OmpR), iron availability (Fur), and oxygen levels (FNR).

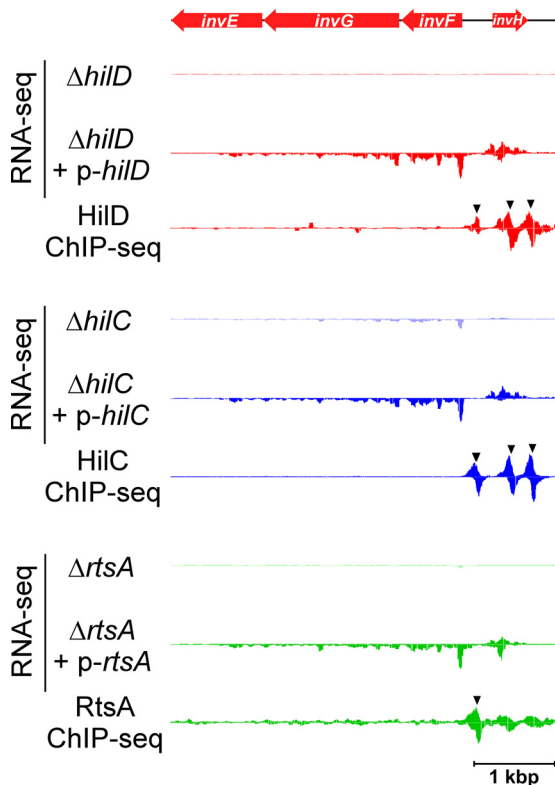
HilD and HilA targets have been identified on a genomic scale using chromatin immunoprecipitation (ChIP) methods and transcription profiling (11, 13). However, some of the regulatory targets for HilA that have been identified using genome-scale approaches are inconsistent with targeted investigations of individual transcripts (18, 24). Moreover, there have been no genome-scale investigations of the regulons of HilC, RtsA, RtsB, InvF, or SprB. Hence, the extent of SPI-1-associated gene regula-

tion is largely unknown. Here, we used a combination of ChIP sequencing (ChIP-seq) and transcriptome sequencing (RNA-seq) to comprehensively map the direct regulatory targets of SPI-1-associated TFs under conditions that mimic the host environment during the invasion process. Thus, we have generated a critical resource for understanding the regulation associated with epithelial cell invasion. Combining ChIP-seq and RNA-seq data allows us to identify all regulated genes and to distinguish direct and indirect regulatory targets. All ChIP-seq and RNA-seq data are freely available for viewing using an intuitive genome browser (<http://www.wadsworth.org/research/scientific-resources/interactive-genomics> or <http://salmonella.wadsworth.org>). Our data reveal a complex web of regulation, with all SPI-1-associated TFs being regulated by at least one other. We also identified many novel regulatory targets for SPI-1-associated TFs. Expression of these genes is therefore tightly coordinated with the invasion process. Notably, many of the novel regulatory targets encode TFs or small RNAs (sRNAs), indicating that the regulatory network associated with invasion extends well beyond the known SPI-1-associated TFs. By comparing gene expression profiles for the regulatory targets of the seven analyzed TFs, we identified 17 genes whose expression closely mirrors that of known SPI-1 genes. These genes are therefore likely to play direct roles in invasion. We also identified clusters of genes with similar expression profiles, suggesting cross talk between invasion gene regulation and other regulatory pathways.

## RESULTS AND DISCUSSION

**Identification of direct regulatory targets for HilD.** HilD is considered the master regulator of *S. Typhimurium* invasion (8, 11). In a previous study, we mapped the binding of HilD across the *S. Typhimurium* genome using ChIP-seq (11). Thus, we identified 11 novel HilD-bound genomic regions. We showed that genes in four of these regions are associated with transcription activation by HilD. However, we assayed transcription activation by HilD in *Escherichia coli*, by fusing regions upstream of candidate genes to the *lacZ* reporter gene; failure to observe transcription activation by HilD in such experiments does not necessarily indicate that HilD does not regulate these genes, since factors specific to *Salmonella* may be required for their activation. To reassess the HilD regulon, we combined ChIP-seq and RNA-seq to map all direct regulatory targets of HilD in *S. Typhimurium* strain 14028s. For ChIP-seq, HilD was C-terminally epitope tagged and expressed from its native locus. For RNA-seq, we compared RNA levels in cells lacking *hilD* to those in cells transiently overexpressing HilD from a plasmid. We defined directly regulated genes as those protein-coding genes for which we observed a significant difference in RNA levels with and without HilD (see Table S1 in the supplemental material) and for which we detected binding by ChIP-seq (see Table S2) within a window extending from 600 bp upstream to 100 bp downstream of the annotated gene/operon start (see Table S3 and Materials and Methods for complete details on identification of directly regulated genes). An example of a directly HilD-regulated transcript is shown in Fig. 1.

Most of the HilD binding events that we detected (see Table S2 in the supplemental material) corresponded to those described in our previous study (11). Nonetheless, our previous study included a few confirmed binding sites that we did not detect in the current study. This is likely due to differences in sensitivity between the sequencing approaches used, consistent with



**FIG 1** Direct regulation of the *invF*-containing transcript by HilD, HilC, and RtsA. Data represent results of RNA-seq and ChIP-seq analysis of HilD, HilC, and RtsA, for the region encompassing *invF* and *invH*. Red arrows represent genes. RNA-seq graphs show sequence read density for one replicate experiment for cells lacking the indicated TF or for cells transiently overexpressing the indicated TF. ChIP-seq graphs show sequence read density for one replicate experiment for the indicated TF. The genes are colored red to indicate positive regulation by the TFs. Black arrowheads indicate ChIP-seq peaks. RNA-seq data are normalized according to the genome position with the 90th percentile for sequence read coverage.

the missed targets all being in regions of low enrichment. We included these previously identified binding sites in our analysis of direct regulatory targets. All previously described direct regulatory targets of HilD were confirmed by our new analysis (Fig. 2). We identified several additional direct regulatory targets: *sinR* (see Fig. S1A), *STM14\_1283* (*STM05010*) (for genes without common names, we list the homologous gene in *S. enterica* serovar Typhimurium strain LT2, assuming that an annotated homologue exists), *STM14\_1613* (*STM1329*), *STM14\_1614* (*STM1330*), *mcpC*, *STM14\_5184* (*STM4310*), *STM14\_5185* (*STM4312*), *STM14\_5186* (*STM4313*), and *STM14\_5189* (Fig. 2). Our previous study identified binding sites for HilD near the start codons of all of these genes (11), but we tested regulation of most using assays of reporter gene fusions in *E. coli* and did not observe transcription activation by HilD. We conclude that these genes require additional, *Salmonella*-specific factors for transcription activation by HilD.

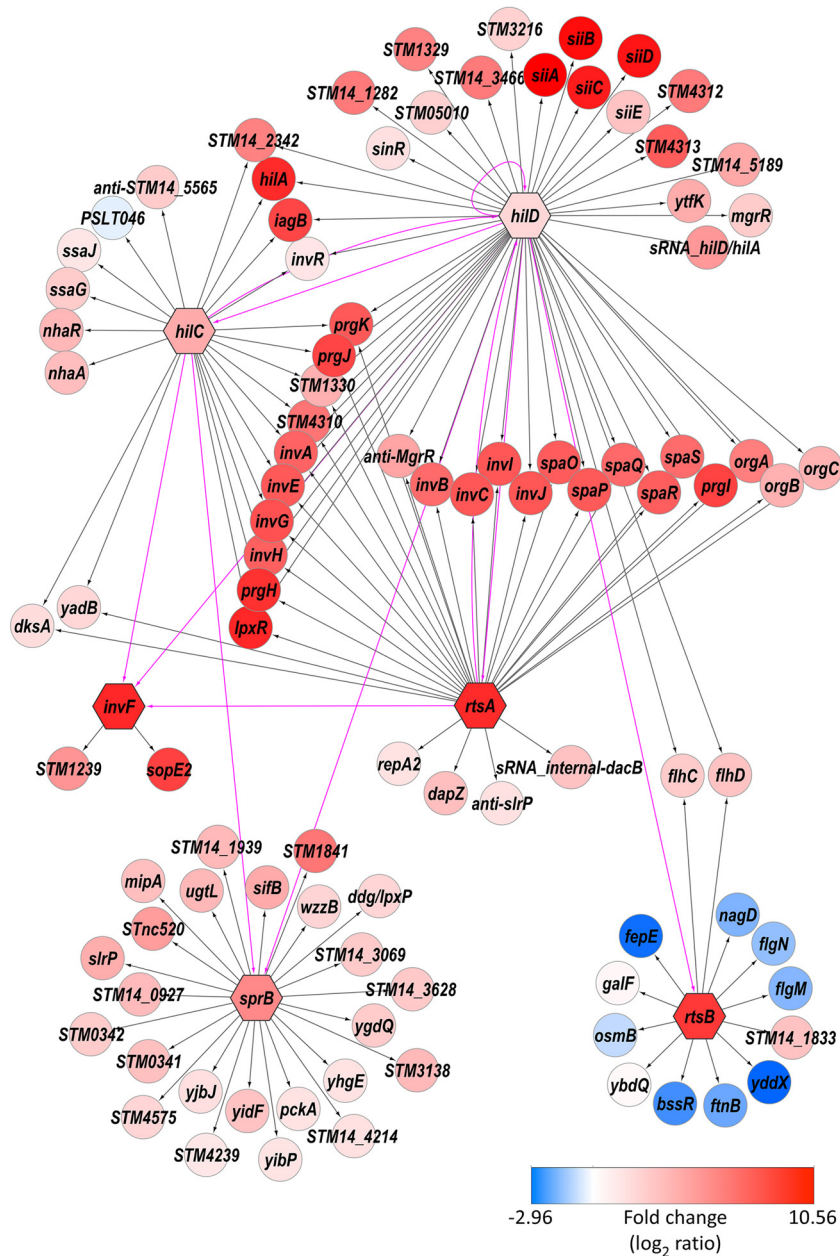
**HilD directly regulates all other SPI-1-associated TFs.** Aside from HilD, four of the SPI-1-associated TFs, HilC, InvF, SprB, and HilA, are encoded within SPI-1. RtsA is a close homologue of HilD and HilC but is encoded outside SPI-1, in an operon that includes another TF-encoding gene, *rtsB*. Both *rtsA* and *rtsB* are strongly

induced under conditions associated with SPI-1 expression (4). Our analysis of the HilD regulon indicated that HilD directly regulates all seven SPI-1-associated TFs, including HilD itself (Fig. 2) (11). This is consistent with the idea that HilD is the apex of a regulatory cascade associated with host cell invasion and that HilD activity is the primary target of environmental signals associated with invasion (9). We detected direct regulation of *sprB* by HilD from a binding site immediately upstream of *sprB*. Visual inspection of the RNA-seq data suggested that *sprB* is also cotranscribed with *hilC*, which is itself a direct HilD target (see Fig. S1B in the supplemental material). Using reverse transcription-PCR (rt-PCR), we detected a transcript that spans the *hilC* and *sprB* genes (see Fig. S2). Given that the level of HilD binding upstream of *hilC* is substantially higher than that upstream of *sprB* (see Fig. S1B), we conclude that HilD regulation of *sprB* occurs predominantly through regulation of the *hilC-sprB* transcript.

**Combining ChIP-seq and RNA-seq to map the entire invasion regulon.** Although a few regulatory targets have been described for SPI-1-associated TFs other than HilD, only the HilD and HilA regulons have been determined on a genomic scale (11, 13). As described above, the ChIP-seq and RNA-seq approaches provide a complementary pipeline for identification of direct regulatory targets. Hence, we used this pipeline to identify the direct regulatory targets of HilC, RtsA, RtsB, InvF, and SprB (see Tables S1 to S3 in the supplemental material). For HilA, we were unable to detect robust ChIP-seq enrichment, suggesting that the epitope tags inactivated HilA function. We therefore limited our analysis of the HilA regulon to RNA-seq (see Table S1).

**The HilD, HilC, and RtsA regulons overlap but are distinct.** HilC and RtsA are close homologues of HilD (7, 10, 11, 23). Previous work has shown that these TFs share many regulatory targets (10, 11, 21, 44). Using ChIP-seq and RNA-seq, we identified 21 direct target genes for HilC (12 transcripts) and 28 for RtsA (9 transcripts) (Fig. 2; see also Table S3 in the supplemental material). As expected, many of the HilD, HilC, and RtsA targets are shared (Fig. 1 and 3A). To determine whether HilD, HilC, and RtsA bind the same DNA sites upstream of shared target genes, we compared the profile of ChIP-seq reads across regions with closely positioned HilD/HilC/RtsA ChIP-seq peaks. The results of this analysis clearly indicate that HilD, HilC, and RtsA bind the same DNA sequences (Fig. 3B). Consistent with this, we were able to infer DNA sequence motifs for HilC (two similar motifs that are almost completely mutually exclusive in their site composition) and RtsA from the ChIP-seq data (Fig. 3C). Although the motifs are similar to each other, subtle differences in these motifs likely explain the overlapping but distinct binding profiles of HilC and RtsA. Although we could not derive a DNA sequence motif for HilD, it is likely that it binds with similar levels of sequence specificity to HilC and RtsA, given the overlap in target sites (Fig. 3B). Our data are consistent with an earlier study showing that HilC and HilD bind overlapping sites upstream of *hilA*, *hilD*, and *hilC* but that HilC binding is less sensitive to mutations within the DNA site (21).

Although the HilC regulatory targets were largely a subset of the HilD targets, HilC binds many more sites across the genome (209 sites for HilC versus 18 sites for HilD, not including the HilD sites from our previous study [11]). However, we detected regulation of nearby genes for only a small fraction of the HilC binding sites (Fig. 3D; see also Fig. S1C in the supplemental material). In contrast, we detected regulation of nearby genes for all but two

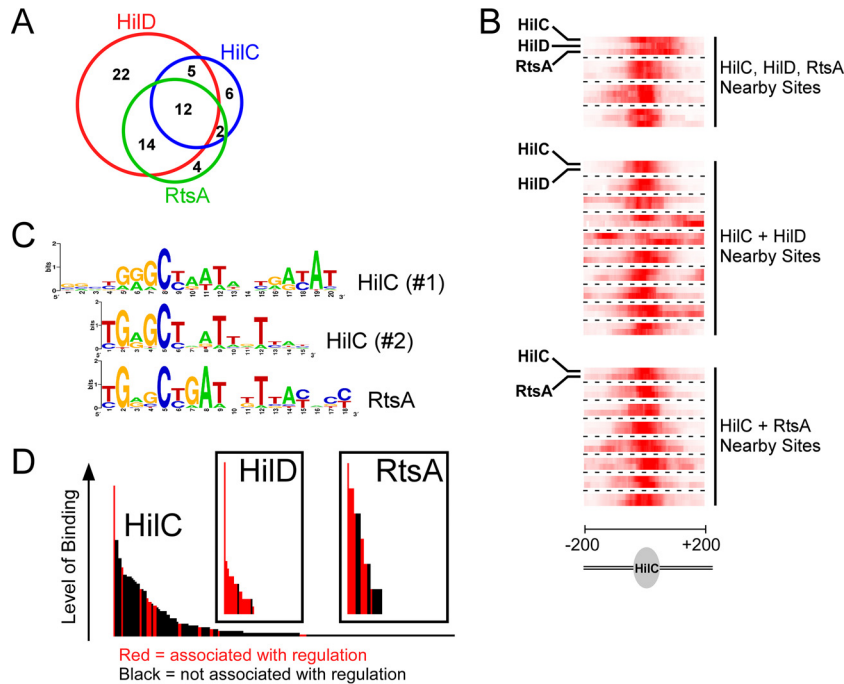


**FIG 2** Regulatory network associated with SPI-1. The hexagon-shaped nodes represent TF-encoding genes *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, and *rtsB*. The circular nodes represent target genes directly regulated by these TFs. The directed edges indicated with pink lines represent the regulatory relationships among the TFs. The directed edges indicated with black lines represent the regulatory relationships between the TFs and their target genes. The node color indicates the log ratio of the differential levels of expression (red = TF activated; blue = TF repressed), and the color scale is shown. Note that the regulatory targets of HilA are omitted because ChIP-seq for this protein was unsuccessful. Where possible, common gene names are shown. For genes without a common name, the name of the orthologous gene from *S. Typhimurium* strain LT2 is shown (“STMxxxx” names). For annotated genes without a common name and without an orthologue in *S. Typhimurium* strain LT2, the *S. Typhimurium* 14028s gene name is shown (“STM14\_xxxx” names). For unannotated genes, we list a brief description of the transcript location (“anti” = antisense to the indicated gene; “sRNA\_internal” = initiation within the indicated gene, in the sense orientation relative to that gene; “sRNA\_hilD/hilA” = initiation in the intergenic region between *hilD* and *hilA*, antisense to the *hilA* 5' UTR).

HilD binding sites and for 50% of the RtsA binding sites (we include regulation of noncoding RNAs, discussed below; Fig. 3D). These data suggest that HilC has a more relaxed DNA sequence specificity than HilD or RtsA, but that HilC is a less potent activator of transcription.

Transcription activation of SPI-1 genes by HilD, HilC, and RtsA is believed to be due to countersilencing of repressive NAPs

H-NS and Hha rather than to direct interaction between the activators and initiating RNA polymerase (29, 30). We compared the locations of HilD, HilC, and RtsA binding sites to the location of H-NS occupancy, as determined by chromatin immunoprecipitation with microarray technology (ChIP-chip) analysis performed using the closely related LT2 strain of *S. Typhimurium* (26). Five of the HilD/HilC/RtsA-bound regions associated with direct tran-



**FIG 3** Overlap of the HilD, HilC, and RtsA regulons. (A) Venn diagram showing the degree of overlap of direct regulatory targets between HilD, HilC, and RtsA. (B) Heat maps showing overlap of ChIP-seq peaks for HilD, HilC, and RtsA. Each block shows data for two or three TFs at a single 400-bp genomic region, centered on a ChIP-seq peak for HilC. The combined, normalized sequence read count from the ChIP-seq data for HilC is shown in the first row of each block, with color intensity representing read density. The one or two additional rows in each block show equivalent data for HilD or RtsA or both, as indicated. (C) Sequence motifs for HilC and RtsA. (D) Association of ChIP-seq peaks with regulation for HilD, HilC, and RtsA. Each of the histograms represents all ChIP-seq peaks (i.e., binding sites for the corresponding TF). Each bar represents a single ChIP-seq peak, and the height of the bar indicates the level of binding (FAT score). Binding sites associated with regulation are indicated by red bars. Binding sites not associated with regulation are indicated by black bars.

scription activation of protein-coding genes outside SPI-1 overlap regions that are likely bound by H-NS (HilD/HilC/RtsA binding sites associated with regulation of *lpxR*, *STM14\_5184*, *siiA*, *mcpC*, and *ssaG*). Thus, many of the target genes of HilD, HilC, and RtsA are likely regulated by a countersilencing mechanism. This is consistent with the fact that *hilD*, *hilC*, and *rtsA* were horizontally acquired by *S. enterica* (45) and that countersilencing is expected to evolve more quickly than other mechanisms of transcription activation (36).

**InvF and SprB broaden the extent of positive regulation associated with invasion.** HilD, HilC, and RtsA all positively regulate *invF*, which is located within SPI-1 and encodes another of the SPI-1-associated TFs (Fig. 1 and 2). Three regulatory targets of InvF, *sopB* (*sigD*), *sicA*, and *sopE*, have been described previously (14, 46) (note that *sopE* is not present in the strain we used but that a close homologue, *sopE2*, is present). *sopB* and *sopE2* encode SPI-1-secreted proteins, and *sicA* encodes a chaperone for SPI-1 secreted proteins. *sopB* and *sopE2* are located outside SPI-1. Our ChIP-seq and RNA-seq data for InvF confirm direct regulation of *sopE2* (Fig. 2; see also Fig. S1D in the supplemental material) and identify *STM14\_1486* (*STM1239*; Fig. 2; see also Fig. S1E) as an additional gene that is positively regulated by InvF. Given that all other described InvF targets are directly involved in the function of the SPI-1-encoded T3SS, our data strongly implicate *STM14\_1486* in this process. Surprisingly, we did not detect InvF binding upstream of *sopB* or *sicA*, despite these genes having been previously shown to be regulated directly by InvF (14). Our RNA-seq data show strong upregulation of *sopB* and *sicA* by InvF (see

Table S1), suggesting that failure to detect binding of InvF by ChIP-seq might be due to partial inactivation of the protein by the epitope tags. This seems unlikely given that we detected robust association of InvF with sites upstream of *sopE2* and *STM14\_1486* (see Fig. S1D and E). An alternative explanation is that InvF binds upstream of *sopB* and *sicA* only when present at high concentrations in the cell, as would be the case in our RNA-seq experiment in which InvF was transiently overexpressed. Since we identified only two regions bound by InvF, we were not able to infer the sequence of its cognate DNA sites. However, a consensus sequence for InvF has been described previously (46), and a similar sequence is found upstream of *sopE2* and *STM14\_1486*, coincident with the InvF-enriched genomic region (see Fig. S3).

HilD and HilC positively regulate *sprB*, most likely cotranscriptionally with *hilC* itself, since *hilC* and *sprB* are cotranscribed (see Fig. S2 in the supplemental material). The only previously described target of SprB is the *siiABCDEF* operon of SPI-4 (24). Our ChIP-seq and RNA-seq data identified 23 direct target genes for SprB, all of which are positively regulated. An example of an SprB-regulated gene, *STM14\_2227* (*STM1841*), is shown in Fig. S1F. Strikingly, we detected no significant regulation of SPI-4 genes (see Table S1), and we detected no binding of SprB within or near SPI-4 (see Table S2 and Fig. S1G). We conclude that SPI-4 is not a regulatory target of SprB. We and others have shown that HilD and HilA positively regulate SPI-4 genes (10, 11, 15, 17), explaining how SPI-4 gene expression is coordinated with that of SPI-1.

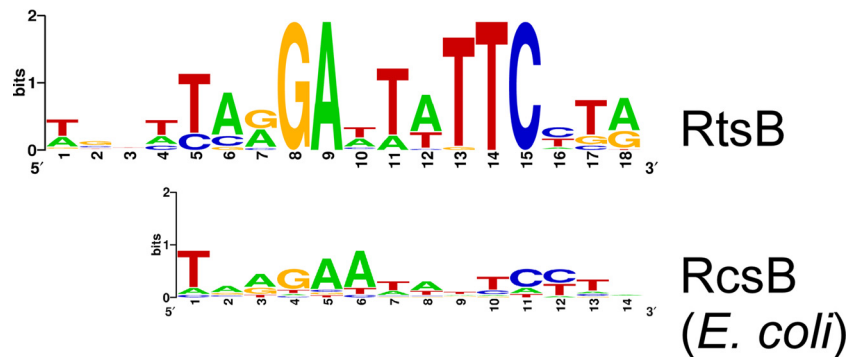


FIG 4 RtsB binding sites that overlap those of RcsB. Data represent results of comparison of the sequence motif for RtsB to the sequence motif for *E. coli* RcsB taken from the PRODORIC database (90).

**RtsB is a negative regulator with a regulon overlapping that of RcsB.** RtsA and RtsB are encoded by genes that are next to one another and are predicted to be cotranscribed in an operon (10). Both *rtsA* and *rtsB* are strongly induced under conditions associated with activation of SPI-1 genes (4). Consistent with this, the *rtsAB* operon is a direct regulatory target of HilD (Fig. 2). Our data also indicate that the two genes downstream of *rtsB* are likely to be cotranscribed with *rtsAB* and that the nearby *STM14\_5184* gene is positively regulated by HilD, HilC, and RtsA (Fig. 2). Thus, although this genomic region falls well outside SPI-1, the genes it contains are coordinately regulated with those in SPI-1. RtsB is a TF that includes a DNA-binding domain belonging to the family of two-component response regulators. However, in contrast to most members of this family, RtsB lacks a receiver domain, suggesting that it cannot directly respond to signaling from a histidine kinase. Our ChIP-seq and RNA-seq data identified 10 direct regulatory targets of RtsB (Fig. 2). With one exception, all these genes are negatively regulated by RtsB. This is in contrast to most two-component system regulators, which predominantly function to activate transcription. We observed negative regulation of the *flhDC* operon that encodes the master regulators of flagellar motility (see Fig. S1H in the supplemental material), consistent with a previous study (10). We also observed negative regulation of *flgM* and *yddX* by RtsB (Fig. 2). *flgM* encodes an anti-sigma factor for FliA ( $\sigma^{28}$ ), the flagellar Sigma factor (47). *yddX* encodes a positive regulator of flagellar synthesis in *E. coli* (48). Thus, RtsB both negatively (through FlhDC and YddX) and positively (indirectly, through FlgM) regulates expression of flagellar motility genes. In contrast, HilD positively regulates FlhDC (see Fig. S1H), suggesting that flagellar regulation during invasion is complex.

Strikingly, the most enriched region in the RtsB ChIP-seq data, as well as five other major ChIP-seq peaks, was upstream of the *std* operon that encodes a fimbrial apparatus whose expression is connected to that of SPI-1 genes (see Fig. S1I in the supplemental material) (49). However, we detected no significant difference in the levels of expression of *std* genes between cells lacking *rtsB* and cells transiently overexpressing RtsB (see Fig. S1I and Table S1). Given the number and strength of the binding sites for RtsB upstream of the *std* operon, we propose that RtsB regulates transcription of these genes under a growth condition different from that used in our study.

As for HilC and RtsA, we were able to infer a DNA motif for RtsB binding by searching for enriched sequences within the RtsB-

bound regions (Fig. 4). This motif closely resembles that of *E. coli* RcsB, a two-component system regulator (Fig. 4). RcsB is known to repress transcription of *flhDC* in both *E. coli* (50) and *S. enterica* (19). Moreover, RcsB regulates *yddX* in *E. coli* (51). Thus, our data suggest regulatory interplay between RcsB and RtsB. Intriguingly, RcsB has been shown previously to regulate genes in the *std* operon (52), further suggesting that RtsB regulates this operon.

**Redefining the HilA regulon using RNA-seq data.** HilA is known to positively regulate the transcripts of *inv* and *spa*, *sic* and *sip*, and *prg* and *org* located within SPI-1 (8), and a combined ChIP-chip and microarray transcription profiling study identified additional HilA targets (13). We failed to detect any binding sites for HilA by ChIP-seq, suggesting that the epitope tags interfered with HilA function. However, RNA-seq data from a  $\Delta$ *hilA* strain with/without transient overexpression of HilA were consistent with regulation of transcripts within SPI-1 (see Table S1 in the supplemental material). Strong regulation of genes outside SPI-1 was observed, but most of these genes are direct regulatory targets of other SPI-1-associated TFs (HilD, HilC, RtsA, InvF, and SprB), and HilA likely regulates expression of one or more of these TFs. Nonetheless, we were able to identify likely direct targets of HilA based on the criteria that (i) the genes were strongly regulated in the RNA-seq experiments and (ii) the genes are not direct regulatory targets of other known SPI-1-associated TFs. Thus, we identified genes in the *pspA-pspD* operon and *pspG* as likely HilA targets (see Fig. S1J). These genes encode phage shock proteins that are essential for virulence, likely because of their role in maintaining proton motive force during intracellular growth (53). We also identified *yiaD* (see Fig. S1K) and *STM14\_1174-8* (*STM1034-8*) as likely direct targets of HilA. Surprisingly, the 13 genes located outside SPI-1 that have been previously reported as direct HilA targets (13) were not significantly regulated in our RNA-seq experiment. One of the previously reported HilA targets for which we did not observe significant regulation is *flhD* (13). We note that two previous studies provided strong evidence that *flhD* is not regulated by HilA (18, 54). We therefore suggest that the previously reported HilA targets outside SPI-1 (13) likely represent false positives.

**Regulation of noncoding RNAs by SPI-1-associated TFs.** Bacterial genomes encode large numbers of small RNAs (sRNAs) that are typically noncoding (55). sRNAs often function as regulators of mRNA stability and/or translation by base pairing with specific mRNA targets (55). Many sRNAs have been identified in

TABLE 1 List of regulated noncoding RNAs identified in this study

Start (bp) <sup>a,b</sup>	Stop (bp) <sup>a,b</sup>	Strand	Description	TF regulator(s)	Fold change <sup>c</sup>
34013*	33811*	–	Antisense to <i>STM14_5565</i>	HilC	2.28 <sup>d</sup>
74905	74971	+	DapZ	RtsA	2.66
868925	868081	–	Antisense to <i>slrP</i>	RtsA	1.25
1343788	1343659	–	STnc520	SprB	4.08
1418433	1418523	+	Overlapping <i>STM14_1613/STM14_1614</i>	HilC	1.18
1603873	1604092	+	MgrR	HilD	2.12
1603938	1603781	–	Overlapping <i>STM14_1833</i>	HilD, RtsA	3.73, 2.05 <sup>e</sup>
3039507	3039347	–	Antisense to <i>hila</i> 5' UTR	HilD	4.23
3065151	3066061	+	InvR	HilD, HilC	1.07, 3.99 <sup>e</sup>
3483808	3483985	+	Within <i>dacB</i>	RtsA	2.54

<sup>a</sup> The numbers represent predicted start/stop genome coordinates for a given noncoding RNA (ncRNA). Numbers marked with an asterisk indicate RNAs encoded on the virulence plasmid.

<sup>b</sup> In some cases, multiple, overlapping fragments of an ncRNA were identified, or overlapping ncRNAs were identified for multiple TFs. These instances were merged into a single unit.

<sup>c</sup> Data represent fold change ( $\log_2$ ) in RNA level for cells showing overexpression of the corresponding TF versus cells in which the TF-encoding gene was deleted.

<sup>d</sup> Data represent averages of fold change values for three overlapping transfrags.

<sup>e</sup> The two values correspond to the 2 TFs, in the order listed in the previous column.

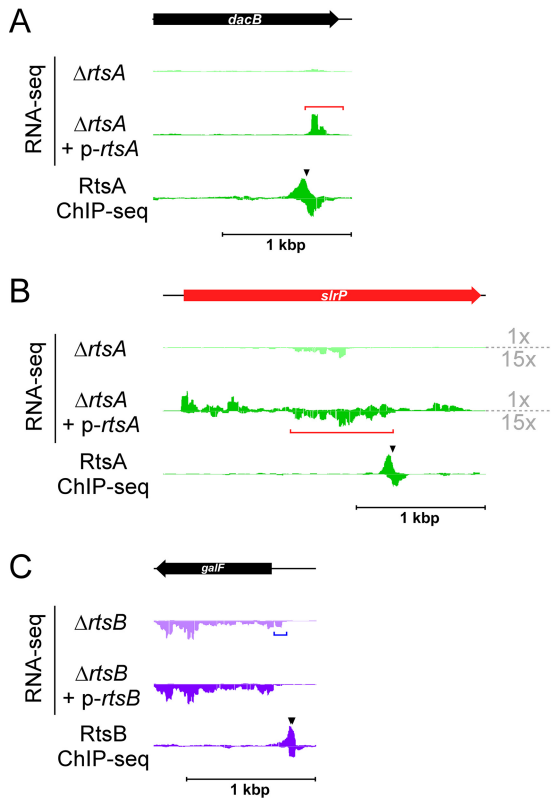
*S. Typhimurium* (56–59). Two sRNAs, InvR and DapZ, have been described previously as direct regulatory targets of HilD (56, 60). We took an unbiased approach to determine whether these and any additional sRNAs are regulated by SPI-1-associated TFs. We applied the Rockhopper RNA-seq analysis tool to our RNA-seq data for each TF, in the absence of any reference genome sequence (61). This method simultaneously assembles transcript fragments (“transfrags”; see Materials and Methods) (62) from the raw RNA-seq data and identifies those transfrags that are significantly differentially expressed between cells lacking a TF and cells transiently overexpressing a TF. Note that a single mRNA can be represented by multiple nonoverlapping transfrags since it is often not possible to assemble a complete transcript sequence from RNA-seq data alone. Regulated transfrags were mapped back to the 14028s genome. We then selected any transfrag whose 5' end fell <200 bp downstream or <100 bp upstream of a TF binding site identified by ChIP-seq. The majority of these directly regulated transfrags corresponded to the 5' portion of mRNAs for protein-coding genes identified in our conventional analysis pipeline. However, we identified several additional directly regulated transfrags. These correspond to known sRNAs, novel sRNAs, and 5' untranslated regions (UTRs) of genes that were not detectably regulated in our conventional analysis. A summary of sRNAs directly regulated by SPI-1-associated TFs is shown in Table 1.

Consistent with a previous study (60), we observed that the InvR sRNA is a direct regulatory target of HilD (Fig. 2). We also observed direct, positive regulation of InvR by HilC (Fig. 2). We did not observe direct regulation of the DapZ sRNA by HilD, in contrast to a previous report (56). Rather, we observed direct regulation of DapZ by RtsA (see Fig. S1L in the supplemental material). Moreover, while we did not detect a HilC-regulated transfrag for DapZ using Rockhopper, we did detect strong binding of HilC upstream of *dapZ* (see Fig. S1L), and visual inspection of the RNA-seq data (see Fig. S1L) strongly supports the idea of direct regulation by HilC. We propose that DapZ is regulated by both RtsA and HilC but not HilD. This model is consistent with a previous study which showed a larger decrease in DapZ levels in cells lacking HilD than in cells lacking either HilC or RtsA (56). Since HilD positively regulates both HilC and RtsA (Fig. 2) (11, 20, 21), loss of HilD

would be expected to result in a larger decrease in DapZ expression than loss of either HilC or RtsA individually.

Few of the sRNAs regulated by SPI-1 TFs (Table 1) have been described previously (59). Four of the regulated sRNAs are transcribed from loci that do not overlap other annotated genes. These include an sRNA that overlaps considerably with the previously described STnc520, whose expression was shown to be strongly induced under conditions associated with activation of SPI-1 (4). This is consistent with our observation that the STnc520-overlapping sRNA is positively regulated by SprB (Fig. 2 and Table 1). We detected regulation of two sRNAs that overlap the 3' ends of protein-coding genes in the sense orientation. One of these is DapZ. The other, an RtsA target, is a novel sRNA that initiates within *dacB* (Fig. 5A). These data are consistent with the recent observation of sense-orientation sequence within protein-coding genes being a rich source of sRNAs (56). We also identified three antisense sRNAs. One of these, a target of HilD and RtsA, is antisense to an sRNA, MgrR, which itself is directly regulated by HilD. A particularly intriguing antisense RNA initiates within *slrP* and is positively regulated by RtsA from a binding site within *slrP* (Fig. 5B). This RNA has been observed previously (4), but regulation has not been described. Strikingly, increased expression of the antisense RNA within *slrP* correlates with increased expression of *slrP* itself; *slrP* mRNA levels are 39-fold higher in cells transiently overexpressing RtsA than in cells lacking *rtsA* (Fig. 5B). Regulation of *slrP* by RtsA has been described previously (10), but the reporter gene fusion used in that study lacked the region within *slrP* that is bound by RtsA. We propose that the previously observed regulation of *slrP* by RtsA (10) occurs indirectly, via SprB (Fig. 2), and that an additional level of regulation by RtsA occurs via the antisense RNA within the *slrP* gene.

**Regulation of transcript variants by SPI-1-associated TFs.** We detected regulated transfrags for several regions that are immediately upstream of protein-coding genes, on the same strand. In some cases, these were disregarded because the upstream gene was identified in our standard RNA-seq analysis as being a direct regulatory target of the same TF. One example of the latter, upstream of *invF*, was notable because the RNA initiates well within the *invH* gene, in the antisense orientation. This transfrag is reg-



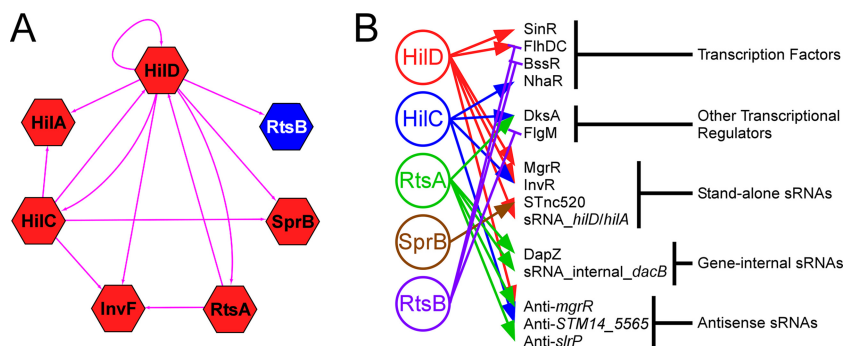
**FIG 5** Identification of regulated sRNAs and 5' UTRs. (A) Data represent results of RNA-seq and ChIP-seq analysis of RtsA, for the region encompassing *dacB*. The sRNA that initiates within *dacB* is indicated by a red bracket. (B) RNA-seq and ChIP-seq data for RtsA, for the region encompassing *slrP*. The sRNA that initiates antisense to *slrP* is indicated by a red bracket. Note that the RNA-seq data for the minus strand are zoomed 15 times relative to the data for the plus strand. (C) RNA-seq and ChIP-seq data for RtsB, for the region encompassing *galF*. The 5' UTR of a *galF* transcript that is regulated by RtsB is indicated by a blue bracket. In all panels, the genes are colored red or black to indicate positive regulation or no regulation by the TFs. Black arrowheads indicate ChIP-seq peaks.

ulated by HilD and HilC. Our RNA-seq data suggested that this RNA is the 5' UTR for the transcript that begins with *invF* (Fig. 1), consistent with earlier studies (12, 63). We confirmed this using rtPCR (see Fig. S4 in the supplemental material). Thus, expression

of the *invF* transcript may also be associated with regulation of *invH* due to the high potential for base pairing between these RNAs. Four other regulated transfrags, three regulated by RtsB and one by HilC, are immediately upstream of genes not significantly regulated according to our standard RNA-seq analysis (see Table S4). An example of such a gene, *galF*, is shown in Fig. 5C. In these cases, we suspected that the genes have multiple transcription start sites, with the most upstream promoter being regulated by the TF in question. Thus, the regulated transcript is a variant with a longer 5' UTR. Published transcription start site mapping supports the idea of multiple transcription start sites for two of the genes, *osmB* and *galF* (4). We confirmed the presence of a long 5' UTR for all four genes by the use of rtPCR (see Fig. S4). Although the overall RNA levels for these genes were not significantly affected by RtsB expression under the conditions that we used (see Table S1), we propose that significant regulation might be observable under other growth conditions, or that the regulated transcript variants might have altered translation efficiency due to their extended 5' UTRs.

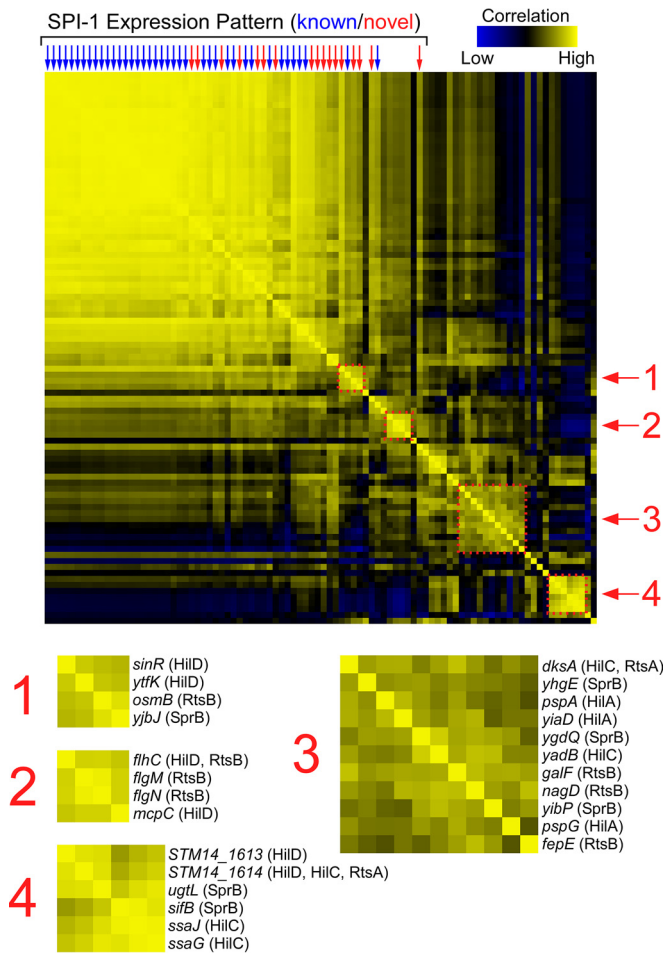
**An expanded regulatory network for invasion.** Cross talk between regulatory pathways is a common theme in bacterial gene regulation (64). Our data demonstrate extensive cross talk in the SPI-1 regulatory network (Fig. 6A). In particular, HilD directly regulates expression of all 7 TFs, including itself. Moreover, as has been described previously, the HilD, HilC, and RtsA regulons overlap extensively (11, 22). In addition to the known SPI-1-associated TFs, our data suggest the involvement of many other regulatory molecules in the gene expression changes that accompany invasion (Fig. 6B). We observed direct regulation of 4 TFs (FlhDC, BssR, SinR, and NhaR) as well as of two other proteins that regulate transcription (FlgM and DksA). We also observed direct regulation of many sRNAs. Some of these sRNAs (e.g., MgrR) are known regulators, whereas others have not been studied previously. Nonetheless, most sRNAs studied to date have been shown to regulate expression of other genes through base-pairing interactions (55). Hence, we propose that *Salmonella* invasion is associated with widespread regulation by sRNAs.

**A core set of invasion genes.** A recent study used RNA-seq to determine expression changes across 22 different growth conditions (4). Using this resource, we selected the expression profiles, for the 22 tested growth conditions, for each of the SPI-1-associated TF target genes. We then grouped the expression pro-



**FIG 6** Extensive cross talk between SPI-1 regulators, and predicted network expansion. (A) The hexagon-shaped nodes represent TF genes *hilD*, *hilC*, *rtsA*, *hilA*, *invF*, *sprB*, and *rtsB*. The directed edges indicated with pink lines represent the regulatory relationships among these TFs. The node color indicates whether the TF is a positive (red) or negative (blue) regulator. (B) Regulation of known and predicted regulators by SPI-1 TFs. Positive regulation is indicated by lines with arrowheads. Negative regulation is indicated by lines with flat heads.





**FIG 7** Hierarchical clustering of gene expression profiles reveals connections between genes regulated by SPI-1-associated TFs. The central heat map shows pairwise correlations between gene expression profiles (data from reference 4) for every direct regulatory target of HilD, HilC, RtsA, InvF, SprB, RtsB, and HilA. Each row/column represents a gene, with genes arrayed in identical orders in rows and columns (note the symmetry of the heat map). Stronger yellow colors indicate higher correlation coefficients, as indicated on the scale (bottom right). Columns indicated by arrows represent genes whose expression profile has an average correlation coefficient of  $>0.5$  for comparisons to all genes located within SPI-1. Blue arrows indicate known invasion genes. Red arrows indicate likely novel invasion genes. The four dashed squares highlight clusters of genes whose expression profiles correlate strongly. These clusters are enlarged and annotated below the heat map.

files using hierarchical clustering. Thus, we were able to identify pairs of genes with similar expression profiles across the 22 growth conditions. Lastly, we compared the expression profiles for each gene by determining correlation coefficients of the 22 expression values for each gene pair. Figure 7 shows the complete set of pairwise correlation coefficients. Each row or column represents a gene identified in our study, and the shading at each row/column intersect indicates how well the expression profiles for the corresponding two genes correlate. Thus, yellow blocks indicate pairs of genes with similar expression levels across the 22 growth conditions analyzed. As expected, known invasion genes formed a tightly clustered group with highly correlated gene expression profiles (Fig. 7). Strikingly, many of the novel regulatory targets identified in our study displayed highly correlated gene expression

profiles with the known invasion genes. Specifically, we identified 17 novel regulatory targets encoded outside SPI-1 whose expression profiles have an average correlation coefficient of  $>0.5$  with the profiles for invasion genes within SPI-1 (Fig. 7 and Table 2). These genes include regulatory targets of multiple TFs, notably, SprB, InvF, and HilA. Given that these 17 genes are direct regulatory targets of SPI-1-associated TFs and that their expression profiles are very similar to those of known invasion genes, we conclude that most or all of these genes are directly involved in the invasion process. Addition of these genes would greatly expand the set of known invasion genes.

**Connecting novel invasion genes with virulence.** Given the importance of the invasion process for infection of host cells, it is perhaps unsurprising that many genes within SPI-1 have been shown to be required for establishing a productive infection in a variety of different model systems, both *in vitro* (i.e., cultured mammalian cells) and *in vivo* (mouse models of infection) (65–68). We analyzed published datasets to determine whether any of the novel invasion genes were identified as being important for productive infection of pigs, chicks, or cows by the use of an oral model of infection (68). Note that other infection models, such as intraperitoneal infection, are less likely to be dependent upon invasion. Of the 17 novel invasion genes, mutants of 14 were analyzed for their ability to infect each of the three hosts. Strikingly, the number of instances where mutants with mutations in these 14 genes were found to be defective in infection was far greater than that expected by chance (binomial test  $P = 4.7E^{-9}$ ). Specifically, mutants with mutations in nine of the novel invasion genes tested (*STM14\_5186*, *STM14\_5184*, *STM14\_5185*, *STM14\_3799* [*STM3138*], *STnc520*, *STM14\_1486*, *nhaA*, *sinR*, and *mcpC*) were defective in establishing an infection in all three infection models (68), and mutants with mutations in *STM14\_2227* were defective in infection of chicks and cows but were not tested in pigs (68). These data suggest that most of the 17 novel invasion genes are directly or indirectly involved in invasion and that their regulation

**TABLE 2** List of putative novel invasion genes

Gene ID <sup>a</sup>	Common name <sup>a</sup>	ST4/74 name <sup>b</sup>	LT2 name <sup>c</sup>	TF <sup>d</sup>
<i>STM14_0358</i>	<i>sinR</i>	<i>SL0300</i>	<i>STM0304</i>	HilD
<i>STM14_5184</i>		<i>SL4247</i>	<i>STM4310</i>	HilD/HilC/RtsA
<i>STM14_5185</i>		<i>SL4248</i>	<i>STM4312</i>	HilD
<i>STM14_5186</i>		<i>SL4249</i>	<i>STM4313</i>	HilD
<i>STM14_0048</i>	<i>nhaA</i>	<i>SL0040</i>	<i>STM0039</i>	HilC
<i>STM14_0049</i>	<i>nhaR</i>	<i>SL0041</i>	<i>STM0040</i>	HilC
<i>STM14_1486</i>		<i>SL1177</i>	<i>STM1239</i>	InvF
<i>STM14_0398</i>		<i>SL0336</i>	<i>STM0341</i>	SprB
<i>STM14_2227</i>		<i>SL1770</i>	<i>STM1841</i>	SprB
<i>STM14_3799</i>		<i>SL3112</i>	<i>STM3138</i>	SprB
<i>STM14_4215</i>	<i>pckA</i>	<i>SL3467</i>	<i>STM3500</i>	SprB
<i>STM14_5097</i>	<i>yjbJ</i>	<i>SL4176</i>	<i>STM4240</i>	SprB
NA	<i>STnc520</i>	<i>STnc520</i>	<i>STnc520</i>	SprB
<i>STM14_1174</i>		<i>SL0973</i>	<i>STM1034</i>	HilA
<i>STM14_1176</i>		<i>SL0975</i>	<i>STM1036</i>	HilA
<i>STM14_1177</i>		<i>SL0976</i>	<i>STM1037</i>	HilA

<sup>a</sup> The genes listed are direct regulatory targets of SPI-1-associated TFs that are encoded outside SPI-1 whose expression profiles (from reference 4) have an average correlation coefficient of  $>0.5$  with the profiles for invasion genes within SPI-1.

<sup>b</sup> Gene name for homologue in strain ST4/74.

<sup>c</sup> Gene name for homologue in strain LT2.

<sup>d</sup> The listed TF(s) directly regulates the corresponding gene.

by SPI-1-associated TFs plays an important role in coordinating invasion processes.

**Regulatory cross talk groups functionally related SPI-1-associated TF target genes.** Hierarchical clustering of gene expression profiles for the SPI-1-associated TF target genes revealed multiple smaller groups of genes with highly correlated expression profiles. Four such clusters are highlighted in Fig. 7. For one of the clusters (cluster 2), the four genes, *flhC*, *flgM*, *flgN*, and *mcpC*, are direct targets of either HilD or RtsB (*flhC* is regulated by both; Fig. 7). Despite being regulated by two different TFs, these genes exhibit similar expression profiles, strongly suggesting an additional, shared regulatory input. For this specific example, the shared regulatory input is known: these genes are all involved in motility and chemotaxis, and their expression is known to be controlled directly or indirectly by the master regulator of flagellar synthesis, FlhDC (69). Given that a connection between SPI-1 and flagellar regulation is well established (19), this cluster validates our approach, indicating that other clusters are very likely to have shared regulatory input that is independent of SPI-1. Cluster 4 (Fig. 7) is particularly interesting because three of the genes in this cluster are associated with secretion by the *Salmonella* pathogenicity island 2 (SPI-2) T3SS that secretes proteins required for intracellular survival and persistence (70): *ssaG* and *ssaJ* encode core components of the SPI-2 T3SS machinery (71, 72), and *sifB* encodes a secreted effector protein (73). *ssaG* and *ssaJ* are regulated by HilC, and *sifB* is regulated by SprB. Thus, our data indicate cross talk between SPI-1 and SPI-2 regulation mediated by HilC and SprB. Moreover, our data indicate a connection between SPI-1 and SPI-2 through three additional genes: *ugtL*, *STM14\_1613*, and *STM14\_1614*. Strikingly, *ugtL*, a target of SprB, is also regulated by SsrB, the master regulator of SPI-1 (74), reinforcing the idea of a role of this gene in cross talk between the two T3SSs. Although expression of SPI-1 and SPI-2 genes is believed to occur at different stages of the infection process, cross talk between the two has been previously described. Specifically, HilD is believed to activate transcription of the genes encoding the SpiR(SsrA)/SsrB two-component regulatory system that is the master regulator of SPI-2 (75). Our data indicate that cross talk between SPI-1 and SPI-2 is more extensive than previously appreciated.

As described above, our hierarchical clustering analysis revealed groups of genes with shared regulatory inputs, which in turn likely indicates shared functions. While it is not immediately clear what the shared regulatory input and shared functions are (e.g., clusters 1 and 3 in Fig. 7), our data will guide future studies of these gene clusters. Moreover, function can be inferred in some cases from the activity of other gene members within a cluster, such as cross talk between SPI-1 and SPI-2 for genes in cluster 4 (Fig. 7). Strikingly, each of the gene clusters includes regulatory targets of different SPI-1-associated TFs. This indicates that cross talk with other regulatory pathways is broadly associated with invasion TF regulation rather than with specific, individual TFs.

**A community resource for *Salmonella* gene regulation.** Our data represent the first global analysis of regulation associated with *Salmonella* invasion. The regulatory network that we describe is highly interconnected, especially in the HilD/HilC/RtsA module, and likely extends well beyond its currently defined gene set due to the regulation of many known and putative regulators by SPI-1 TFs. Analysis of gene expression profiles for genes within the

network has revealed a large set of genes that are likely to be novel invasion genes, and facilitated identification of functionally related genes based on their shared expression profiles. We anticipate that our data will provide a valuable resource to other groups studying *Salmonella* virulence. Hence, we have made our data freely accessible at <http://www.wadsworth.org/research/scientific-resources/interactive-genomics> or <http://salmonella.wadsworth.org> using JBrowse, a web-based browser for visualizing genome-scale datasets (76).

## MATERIALS AND METHODS

**Strains and plasmids.** All strains and plasmids used in this work are listed in Table S5 in the supplemental material. All oligonucleotides used in this work are listed in Table S6. DNA sequence manipulation was performed using BioWord (77). Genomic analyses for 14028s used NCBI reference sequence NC\_016856.1.

All strains of *Salmonella enterica* subspecies *enterica* serovar Typhimurium used in this work were 14028s strains (78). Three C-terminal FLAG tags were inserted chromosomally for *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, *rtsB*, and *hilA*, using the FRUIT method, as previously described (79), to yield strains AMD475, AMD474, AMD476, AMD478, AMD508, AMD477, and AMD473, respectively. Deletions of *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, *rtsB*, and *hilA* were constructed by P22 transduction from existing deletion strains to yield strains CDS024, CDS022, CDS020, CDS028, CDS026, CDS030, and CDS032, respectively. Both the donor and recipient strains were 14028s or derivatives of 14028s. Note that we chose to recreate these mutations by P22 transduction to eliminate the possibility that the strains contained any other differences from one another or from the parent wild-type strain.

Plasmids pBLP013, pBLP011, and pBLP010 have been described previously (11) and are derivatives of pBAD24 (80) that express *hilD*, *hilC*, and *rtsA*, respectively, from the P<sub>BAD</sub> promoter. pCDS001, pCDS002, pCDS003, and pCDS004 are derivatives of pBAD24 (80) that express *rtsB*, *sprB*, *invF*, and *hilA*, respectively, from the P<sub>BAD</sub> promoter. These plasmids were constructed by amplifying the corresponding genes from strain 14028s by colony PCR using oligonucleotides listed in Table S6 in the supplemental material, followed by insertion into pBAD24 cut with NcoI using an In-Fusion kit (Clontech).

**ChIP-seq.** Strains AMD475, AMD474, AMD476, AMD478, AMD508, and AMD477 (14028s with C-terminally FLAG-tagged *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, and *rtsB*, respectively) were used for ChIP-seq. Cells were grown overnight in LB (0.17 M NaCl), subcultured 1 in 50 in high-salt LB (0.3 M NaCl), and shaken at 50 rpm at 37°C for 5.5 h. ChIP-seq was performed using M2 anti-FLAG antibody (Sigma), as described previously (81). Two biological replicates were performed for each ChIP-seq experiment.

**ChIP-seq peak identification.** Identification of peaks from the ChIP-seq data was performed as described previously (81), except that the shift between forward and reverse strands was assumed to be at least 20 nucleotides (20 nt). All ChIP-seq peaks are listed in Table S2 in the supplemental material.

**Motif identification and scoring.** To identify enriched sequence motifs from ChIP-seq data, we first extracted 100-bp sequences centered on each of the ChIP-seq peaks. In cases where these sequences overlapped, we merged them into a single sequence of >100 bp. We then used MEME-ChIP (with the default parameters, except that the “any number of sequences” option was selected) to identify enriched sequence motifs. Only motifs with an *E* value of <0.01 are reported.

**RNA-seq.** Strains CDS024, CDS022, CDS020, CDS028, CDS026, CDS030, and CDS032 (14028s strains with an unmarked deletion of *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, *rtsB*, and *hilA*, respectively) were used for RNA-seq. These strains were transformed either with empty pBAD24 or with pBAD24 derivatives pBLP013, pBLP011, pBLP010, pCDS003, pCDS002, pCDS001, and pCDS004, which express *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, *rtsB*, and *hilA*, respectively (the pBAD24-encoded TF matched the deleted TF

gene for each strain). Cells were grown overnight in LB (0.17 M NaCl), subcultured 1 in 50 in high-salt LB (0.3 M NaCl), and shaken at 50 rpm at 37°C for 5.5 h, and arabinose was added (0.2% final concentration) for 7 min before harvesting was performed. Note that 7 min was chosen as the induction time because a previous study had showed that expression of a regulator leads to direct regulatory effects that occur between 5 and 10 min after induction (82) and because transient induction of TF expression limits indirect regulatory effects. RNA purification, DNase I treatment, rRNA removal, and Illumina library construction were performed as described previously (83). Two biological replicates were performed for each RNA-seq experiment.

**Identification of direct regulatory targets of SPI-1-associated TFs using genome sequence and annotation.** The following procedure was applied for each of HilD, HilC, RtsA, InvF, SprB, and RtsB. RNA-seq data were analyzed using Rockhopper (version 2.02, default parameters) (84), comparing expression of TF-deleted cells with results determined using empty vector and cells transiently overexpressing the TF. Based on this analysis, annotated genes were selected if their RNA levels were significantly different ( $q = <0.01$ ) by  $>2$ -fold in cells lacking the TF and cells in which the TF was transiently expressed. All RNA-seq analysis data generated using a reference genome are listed in Table S1 in the supplemental material. We then compared the positions of the starts of these regulated genes to the positions of ChIP-seq peaks for the corresponding TF. Significantly regulated genes were classified as “direct regulatory targets” if the start codon of the gene was determined to be  $\leq 600$  bp downstream or  $\leq 100$  bp upstream of a ChIP-seq peak. All direct regulatory targets, and associated metadata, are listed in Table S3.

**Network analysis.** The gene regulatory network graph was constructed and visualized using Cytoscape software (version 3.1.1) (85). TF genes *hilD*, *hilC*, *rtsA*, *invF*, *sprB*, and *rtsB* were assigned as source nodes, and their target genes were assigned as target nodes. The regulatory relationships between the TFs and their target genes were represented by directed edges. Node identity is double-encoded as shape, indicating source or target nodes, and color, indicating the level of regulation. For target genes regulated by multiple TFs, the level of regulation shown is for the highest TF in the hierarchy HilD  $>$  HilC  $>$  RtsA  $>$  InvF  $>$  RtsB  $>$  SprB.

**Identification of direct regulatory targets of SPI-1-associated TFs without genome sequence or annotation.** The following procedure was applied for each of HilD, HilC, RtsA, InvF, SprB, and RtsB. We used Rockhopper to analyze RNA-seq data without a reference genome sequence or annotation (version 2.02; default parameters were used except for omission of a reference genome) (61). We then selected any transcript fragments (“transfrags”) (62) whose RNA level was significantly different ( $q = <0.01$ ) by  $>2$ -fold in cells lacking the TF and cells in which the TF was transiently expressed. We aligned these RNA sequences to the 14028s genome using CLC Genomics Workbench (version 6; default parameters). We then compared the positions of the starts of these regulated transfrags to the positions of ChIP-seq peaks for the corresponding TF. Significantly regulated transfrags were classified as direct regulatory targets if the 5' end of the transfrag was determined to be  $\leq 200$  bp downstream or  $\leq 100$  bp upstream of a ChIP-seq peak.

**Analysis of overlapping HilD, HilC, and RtsA binding sites.** We selected all HilD or RtsA ChIP-seq peaks that were located within 100 bp of a HilC ChIP-seq peak. If multiple HilC peaks were found within 100 bp of a HilD/RtsA peak, we selected the HilC peak with the highest Fold Above Threshold (FAT) score. For 400 bp surrounding each of the HilC peaks, we determined the sequence read coverage on each strand for each of the two replicate ChIP-seq datasets for HilC and for HilD/RtsA (four ChIP-seq peaks were common to all 3 TFs). We normalized the sequence read coverage in each case to the highest value in the 400-bp window. We then combined data for both strands and for both replicates for each TF. These data are represented by the heat maps in Fig. 3B. Heat maps were generated using Java TreeView (contrast set to 3.0) (86).

**Hierarchical clustering analysis.** For all direct regulatory targets of SPI-1-associated TFs (see Table S3 in the supplemental material), as well as predicted direct targets of HilA, we identified annotated homologues in strain SL4/74 using BLAST (87) (the best match was selected unless no matches scored better than an  $E$  value of  $1E^{-10}$ ). We then extracted expression profiles for these genes for the 29 RNA-seq experiments described in reference 4. We performed hierarchical clustering of the data using Cluster (version 3.0; “uncentered Pearson correlation,” “average linkage” settings) (88, 89). We then determined Pearson correlation coefficients for all pairwise comparisons of expression profiles. These correlation coefficients were displayed as a heat map using Java TreeView (86).

**rtPCR.** For analysis of *hilC* and *sprB*, 14028s cells were grown overnight in LB (0.17 M NaCl), subcultured 1 in 20 in high-salt LB (0.3 mM NaCl), and shaken at 50 rpm at 37°C to achieve an optical density at 600 nm ( $OD_{600}$ ) of 0.8 to 0.9, and arabinose was added (0.2% final concentration) for 7 min before cells were harvested. For analysis of *invF* 5' UTR, RNA was taken from RNA-seq samples for CDS024 plus pBLP013. For analysis of *STM14\_5569* 5' UTR, RNA was taken from RNA-seq samples for CDS022 plus pBLP011. For analysis of *ybdQ*, *osmB*, and *galF* 5' UTRs, RNA was taken from RNA-seq samples for CDS030 plus pCDS001. RNA was purified and treated with DNase I as previously described (83). RNA (1 to 2  $\mu$ g) was reverse transcribed using Superscript III with 2  $\mu$ l 100  $\mu$ M random hexamer, following the manufacturer's instructions. A control sample was treated identically but without addition of reverse transcriptase. Following reverse transcription, samples were diluted to 200  $\mu$ l. cDNA was PCR amplified and analyzed by agarose gel electrophoresis. Oligonucleotides used for PCR are listed in Table S6 in the supplemental material.

**Accession numbers.** All raw RNA-seq and ChIP-seq data (.fastq file format) are available at the EBI ArrayExpress database with accession numbers E-MTAB-3987 and E-MTAB-3765.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01024-16/-/DCSupplemental>.

Figure S1, PDF file, 0.4 MB.  
Figure S2, PDF file, 0.1 MB.  
Figure S3, PDF file, 0.1 MB.  
Figure S4, PDF file, 0.1 MB.  
Table S1, XLSX file, 2.8 MB.  
Table S2, XLSX file, 0.02 MB.  
Table S3, DOCX file, 0.03 MB.  
Table S4, DOCX file, 0.01 MB.  
Table S5, DOCX file, 0.02 MB.  
Table S6, DOCX file, 0.02 MB.

## ACKNOWLEDGMENTS

We thank David Grainger, Keith Derbyshire, and Todd Gray for helpful discussion and comments on the manuscript. We thank Fred Heffron for providing strains. We thank BEI Resources for distributing strains. We thank the Wadsworth Center's Applied Genomic Technologies Core Facility for Sanger sequencing, the Wadsworth Center's Bioinformatics and Statistics Core Facility for help setting up the JBrowse server, and the Wadsworth Center's Media and Tissue Culture Core Facility for media services. We thank the University at Buffalo Next Generation Sequencing Core facility for Illumina sequencing.

This work was supported by Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), Department of Health and Human Services, under contract no. HHSN272200900040C, and through the NIH Director's New Innovator Award Program (1DP2OD007188; J.T.W.). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## FUNDING INFORMATION

This work, including the efforts of Joseph Thomas Wade, was funded by HHS | National Institutes of Health (NIH) (1DP2OD007188). This work, including the efforts of Chunhong Mao, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (HHSN272200900040C).

## REFERENCES

- Hohmann EL. 2001. Nontyphoidal salmonellosis. *Clin Infect Dis* 32: 263–269. <http://dx.doi.org/10.1086/318457>.
- Hendriksen RS, Vieira AR, Karlsmose S, Lo Fo Wong DM, Jensen AB, Wegener HC, Aarestrup FM. 2011. Global monitoring of *Salmonella* serovar distribution from the World Health Organization global Food-borne Infections Network Country Databank: results of quality assured laboratories from 2001 to 2007. *Foodborne Pathog Dis* 8:887–900. <http://dx.doi.org/10.1089/fpd.2010.0787>.
- Lostro CP, Lee CA. 2001. The *Salmonella* pathogenicity island-1 type III secretion system. *Microbes Infect* 3:1281–1291. [http://dx.doi.org/10.1016/S1286-4579\(01\)01488-5](http://dx.doi.org/10.1016/S1286-4579(01)01488-5).
- Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL, Canals R, Grissom JE, Conway T, Hokamp K, Hinton JC. 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. *Cell Host Microbe* 14:683–695. <http://dx.doi.org/10.1016/j.chom.2013.11.010>.
- Bajaj V, Lucas RL, Hwang C, Lee CA. 1996. Co-ordinate regulation of *Salmonella* Typhimurium invasion genes by environmental and regulatory factors is mediated by control of *hilA* expression. *Mol Microbiol* 22:703–714. <http://dx.doi.org/10.1046/j.1365-2958.1996.d01-1718.x>.
- Lee CA, Falkow S. 1990. The ability of *Salmonella* to enter mammalian cells is affected by bacterial growth state. *Proc Natl Acad Sci U S A* 87: 4304–4308. <http://dx.doi.org/10.1073/pnas.87.11.4304>.
- Schechter LM, Damrauer SM, Lee CA. 1999. Two AraC/XylS family members can independently counteract the effect of repressing sequences upstream of the *hilA* promoter. *Mol Microbiol* 32:629–642. <http://dx.doi.org/10.1046/j.1365-2958.1999.01381.x>.
- Ellermeier JR, Slauch JM. 2007. Adaptation to the host environment: regulation of the SPI1 type III secretion system in *Salmonella enterica* serovar Typhimurium. *Curr Opin Microbiol* 10:24–29. <http://dx.doi.org/10.1016/j.mib.2006.12.002>.
- Golubeva YA, Sadik AY, Ellermeier JR, Slauch JM. 2012. Integrating global regulatory input into the *Salmonella* pathogenicity island 1 type III secretion system. *Genetics* 190:79–90. <http://dx.doi.org/10.1534/genetics.111.132779>.
- Ellermeier CD, Slauch JM. 2003. RtsA and RtsB coordinately regulate expression of the invasion and flagellar genes in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 185:5096–5108. <http://dx.doi.org/10.1128/JB.185.17.5096-5108.2003>.
- Petrone BL, Stringer AM, Wade JT. 2014. Identification of HilD-regulated genes in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 196:1094–1101. <http://dx.doi.org/10.1128/JB.01449-13>.
- Akbar S, Schechter LM, Lostro CP, Lee CA. 2003. AraC/XylS family members, HilD and HilC, directly activate virulence gene expression independently of HilA in *Salmonella* Typhimurium. *Mol Microbiol* 47: 715–728. <http://dx.doi.org/10.1046/j.1365-2958.2003.03322.x>.
- Thijs IM, De Keersmaecker SC, Fadda A, Engelen K, Zhao H, McClelland M, Marchal K, Vanderleyden J. 2007. Delineation of the *Salmonella enterica* serovar Typhimurium HilA regulon through genome-wide location and transcript analysis. *J Bacteriol* 189:4587–4596. <http://dx.doi.org/10.1128/JB.00178-07>.
- Darwin KH, Miller VL. 1999. InvF is required for expression of genes encoding proteins secreted by the SPI1 type III secretion apparatus in *Salmonella* Typhimurium. *J Bacteriol* 181:4949–4954.
- De Keersmaecker SC, Marchal K, Verhoeven TL, Engelen K, Vanderleyden J, Detweiler CS. 2005. Microarray analysis and motif detection reveal new targets of the *Salmonella enterica* serovar Typhimurium HilA regulatory protein, including *hilA* itself. *J Bacteriol* 187:4381–4391. <http://dx.doi.org/10.1128/JB.187.13.4381-4391.2005>.
- Eichelberg K, Galán JE. 1999. Differential regulation of *Salmonella typhimurium* type III secreted proteins by pathogenicity island 1 (SPI-1)-encoded transcriptional activators InvF and hilA. *Infect Immun* 67: 4099–4105.
- Main-Hester KL, Colpitts KM, Thomas GA, Fang FC, Libby SJ. 2008. Coordinate regulation of *Salmonella* pathogenicity island 1 (SPI1) and SPI4 in *Salmonella enterica* serovar Typhimurium. *Infect Immun* 76: 1024–1035. <http://dx.doi.org/10.1128/IAI.01224-07>.
- Singer HM, Kühne C, Deditius JA, Hughes KT, Erhardt M. 2014. The *Salmonella* Spi1 virulence regulatory protein HilD directly activates transcription of the flagellar master operon *flhDC*. *J Bacteriol* 196:1448–1457. <http://dx.doi.org/10.1128/JB.01438-13>.
- Mousslim C, Hughes KT. 2014. The effect of cell growth phase on the regulatory cross-talk between flagellar and Spi1 virulence gene expression. *PLoS Pathog* 10:e1003987. <http://dx.doi.org/10.1371/journal.ppat.1003987>.
- Ellermeier CD, Ellermeier JR, Slauch JM. 2005. HilD, HilC and RtsA constitute a feed forward loop that controls expression of the SPI1 type three secretion system regulator *hilA* in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* 57:691–705. <http://dx.doi.org/10.1111/j.1365-2958.2005.04737.x>.
- Olekhovich IN, Kadner RJ. 2002. DNA-binding activities of the HilC and HilD virulence regulatory proteins of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 184:4148–4160. <http://dx.doi.org/10.1128/JB.184.15.4148-4160.2002>.
- Saini S, Ellermeier JR, Slauch JM, Rao CV. 2010. The role of coupled positive feedback in the expression of the SPI1 type three secretion system in *Salmonella*. *PLoS Pathog* 6:e1001025. <http://dx.doi.org/10.1371/journal.ppat.1001025>.
- Eichelberg K, Hardt WD, Galán JE. 1999. Characterization of SprA, an AraC-like transcriptional regulator encoded within the *Salmonella typhimurium* pathogenicity island 1. *Mol Microbiol* 33:139–152. <http://dx.doi.org/10.1046/j.1365-2958.1999.01458.x>.
- Saini S, Rao CV. 2010. SprB is the molecular link between *Salmonella* pathogenicity island 1 (SPI1) and SPI4. *J Bacteriol* 192:2459–2462. <http://dx.doi.org/10.1128/JB.00047-10>.
- Dillon SC, Dorman CJ. 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* 8:185–195. <http://dx.doi.org/10.1038/nrmicro2261>.
- Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, Hinton JC. 2006. H-NS mediates the silencing of laterally acquired genes in Bacteria. *PLoS Pathog* 2:e81. <http://dx.doi.org/10.1371/journal.ppat.0020081>.
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC. 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313:236–238. <http://dx.doi.org/10.1126/science.1128794>.
- Schechter LM, Jain S, Akbar S, Lee CA. 2003. The small nucleoid-binding proteins H-NS, HU, and Fis affect *hilA* expression in *Salmonella enterica* serovar Typhimurium. *Infect Immun* 71:5432–5435. <http://dx.doi.org/10.1128/IAI.71.9.5432-5435.2003>.
- Olekhovich IN, Kadner RJ. 2006. Crucial roles of both flanking sequences in silencing of the *hilA* promoter in *Salmonella enterica*. *J Mol Biol* 357:373–386. <http://dx.doi.org/10.1016/j.jmb.2006.01.007>.
- Olekhovich IN, Kadner RJ. 2007. Role of nucleoid-associated proteins Hha and H-NS in expression of *Salmonella enterica* activators HilD, HilC, and RtsA required for cell invasion. *J Bacteriol* 189:6882–6890. <http://dx.doi.org/10.1128/JB.00905-07>.
- Queiroz MH, Madrid C, Paytubi S, Balsalobre C, Juárez A. 2011. Integration host factor alleviates H-NS silencing of the *Salmonella enterica* serovar Typhimurium master regulator of SPI1, *hilA*. *Microbiology* 157: 2504–2514. <http://dx.doi.org/10.1099/mic.0.049197-0>.
- Mangan MW, Lucchini S, Ó Cróinín T, Fitzgerald S, Hinton JC, Dorman CJ. 2011. Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar Typhimurium. *Microbiology* 157:1075–1087. <http://dx.doi.org/10.1099/mic.0.046359-0>.
- Wilson RL, Libby SJ, Freet AM, Boddicker JD, Fahlen TF, Jones BD. 2001. Fis, a DNA nucleoid-associated protein, is involved in *Salmonella typhimurium* SPI-1 invasion gene expression. *Mol Microbiol* 39:79–88. <http://dx.doi.org/10.1046/j.1365-2958.2001.02192.x>.
- Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JC, Dorman CJ. 2004. A global role for Fis in the transcriptional control of metabolism and type III secretion in *Salmonella enterica* serovar Typhimurium. *Microbiology* 150:2037–2053. <http://dx.doi.org/10.1099/mic.0.27209-0>.
- Cameron AD, Dorman CJ. 2012. A fundamental regulatory mechanism operating through *OmpR* and DNA topology controls expression of *Salmonella* pathogenicity islands SPI-1 and SPI-2. *PLoS Genet* 8:e1002615. <http://dx.doi.org/10.1371/journal.pgen.1002615>.

36. Will WR, Bale DH, Reid PJ, Libby SJ, Fang FC. 2014. Evolutionary expansion of a regulatory network by counter-silencing. *Nat Commun* 5:5270.
37. Teixidó L, Carrasco B, Alonso JC, Barbé J, Campoy S. 2011. Fur activates the expression of *Salmonella enterica* pathogenicity island 1 by directly interacting with the *hilD* operator *in vivo* and *in vitro*. *PLoS One* 6:e19711. <http://dx.doi.org/10.1371/journal.pone.0019711>.
38. Ellermeier JR, Slauch JM. 2008. Fur regulates expression of the *Salmonella* pathogenicity island 1 type III secretion system through HilD. *J Bacteriol* 190:476–486. <http://dx.doi.org/10.1128/JB.00926-07>.
39. Martínez LC, Yakhnin H, Camacho MI, Georgellis D, Babitzke P, Puente JL, Bustamante VH. 2011. Integration of a complex regulatory cascade involving the SirA/BarA and Csr global regulatory systems that controls expression of the *Salmonella* SPI-1 and SPI-2 virulence regulons through HilD. *Mol Microbiol* 80:1637–1656. <http://dx.doi.org/10.1111/j.1365-2958.2011.07674.x>.
40. Chubiz JE, Golubeva YA, Lin D, Miller LD, Slauch JM. 2010. FlhZ regulates expression of the *Salmonella* pathogenicity island 1 invasion locus by controlling HilD protein activity in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 192:6261–6270. <http://dx.doi.org/10.1128/JB.00635-10>.
41. Baxter MA, Fahlen TF, Wilson RL, Jones BD. 2003. HilE interacts with HilD and negatively regulates *hilA* transcription and expression of the *Salmonella enterica* serovar Typhimurium invasive phenotype. *Infect Immun* 71:1295–1305. <http://dx.doi.org/10.1128/IAI.71.3.1295-1305.2003>.
42. Quinn HJ, Cameron AD, Dorman CJ. 2014. Bacterial regulon evolution: distinct responses and roles for the identical OmpR proteins of *Salmonella* Typhimurium and *Escherichia coli* in the acid stress response. *PLoS Genet* 10:e1004215. <http://dx.doi.org/10.1371/journal.pgen.1004215>.
43. Van Immerseel F, Eeckhaut V, Boyen F, Pasmans F, Haesebrouck F, Ducatelle R. 2008. Mutations influencing expression of the *Salmonella enterica* serovar Enteritidis pathogenicity island I key regulator *hilA*. *Antonie Van Leeuwenhoek* 94:455–461. <http://dx.doi.org/10.1007/s10482-008-9262-y>.
44. Schechter LM, Lee CA. 2001. AraC/XylS family members, HilC and HilD, directly bind and derepress the *Salmonella* Typhimurium *hilA* promoter. *Mol Microbiol* 40:1289–1299. <http://dx.doi.org/10.1046/j.1365-2958.2001.02462.x>.
45. Mirold S, Ehrbar K, Weissmüller A, Prager R, Tschäpe H, Rüssmann H, Hardt WD. 2001. *Salmonella* host cell invasion emerged by acquisition of a mosaic of separate genetic elements, including *Salmonella* pathogenicity island 1 (SPI1), SPI5, and *sopE2*. *J Bacteriol* 183:2348–2358. <http://dx.doi.org/10.1128/JB.183.7.2348-2358.2001>.
46. Darwin KH, Miller VL. 2001. Type III secretion chaperone-dependent regulation: activation of virulence genes by SicA and InvF in *Salmonella typhimurium*. *EMBO J* 20:1850–1862. <http://dx.doi.org/10.1093/emboj/20.8.1850>.
47. Ohnishi K, Kutsukake K, Suzuki H, Lino T. 1992. A novel transcriptional regulation mechanism in the flagellar regulon of *Salmonella typhimurium*: an antisigma factor inhibits the activity of the flagellum-specific sigma factor, sigma F. *Mol Microbiol* 6:3149–3157. <http://dx.doi.org/10.1111/j.1365-2958.1992.tb01771.x>.
48. Kim JS, Kim YJ, Seo S, Seong MJ, Lee K. 2015. Functional role of *bdm* during flagella biogenesis in *Escherichia coli*. *Curr Microbiol* 70:369–373. <http://dx.doi.org/10.1007/s00284-014-0729-y>.
49. López-Garrido J, Casadesús J. 2012. Crosstalk between virulence loci: regulation of *Salmonella enterica* pathogenicity island 1 (SPI-1) by products of the *std* fimbrial operon. *PLoS One* 7:e30499. <http://dx.doi.org/10.1371/journal.pone.0030499>.
50. Francez-Charlot A, Laugel B, Van Gemert A, Dubarry N, Wiorowski F, Castanié-Cornet MP, Gutierrez C, Cam K. 2003. RcsCDB His–Asp phosphorelay system negatively regulates the *flhDC* operon in *Escherichia coli*. *Mol Microbiol* 49:823–832. <http://dx.doi.org/10.1046/j.1365-2958.2003.03601.x>.
51. Francez-Charlot A, Castanié-Cornet MP, Gutierrez C, Cam K. 2005. Osmotic regulation of the *Escherichia coli* *bdm* (biofilm-dependent modulation) gene by the RcsCDB His–Asp phosphorelay. *J Bacteriol* 187:3873–3877. <http://dx.doi.org/10.1128/JB.187.11.3873-3877.2005>.
52. Wang Q, Zhao Y, McClelland M, Harshey RM. 2007. The RcsCDB signaling system and swarming motility in *Salmonella enterica* serovar Typhimurium: dual regulation of flagellar and SPI-2 virulence genes. *J Bacteriol* 189:8447–8457. <http://dx.doi.org/10.1128/JB.01198-07>.
53. Darwin AJ. 2013. Stress relief during host infection: the phage shock protein response supports bacterial virulence in various ways. *PLoS Pathog* 9:e1003388. <http://dx.doi.org/10.1371/journal.ppat.1003388>.
54. Teplitski M, Goodier RI, Ahmer BM. 2003. Pathways leading from BarA/SirA to motility and virulence gene expression in *Salmonella*. *J Bacteriol* 185:7257–7265. <http://dx.doi.org/10.1128/JB.185.24.7257-7265.2003>.
55. Waters LS, Storz G. 2009. Regulatory RNAs in bacteria. *Cell* 136:615–628. <http://dx.doi.org/10.1016/j.cell.2009.01.043>.
56. Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* 31:4005–4019. <http://dx.doi.org/10.1038/emboj.2012.229>.
57. Padalón-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, Margalit H, Altuvia S. 2008. Small RNAs encoded within genetic islands of *Salmonella* Typhimurium show host-induced expression and role in virulence. *Nucleic Acids Res* 36:1913–1927. <http://dx.doi.org/10.1093/nar/gkn050>.
58. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JC, Vogel J. 2008. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* 4:e1000163. <http://dx.doi.org/10.1371/journal.pgen.1000163>.
59. Kröger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hébrard M, Händler K, Colgan A, Leekitcharoenphon P, Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ, Thomson NR, Vogel J, Hinton JC. 2012. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci U S A* 109:E1277–E1286. <http://dx.doi.org/10.1073/pnas.1201061109>.
60. Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, Vogel J. 2007. A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol Microbiol* 66:1174–1191. <http://dx.doi.org/10.1111/j.1365-2958.2007.05991.x>.
61. Tjaden B. 2015. *De novo* assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* 16:1. <http://dx.doi.org/10.1186/s13059-014-0572-2>.
62. Kampa D, Cheng J, Kapranov P, Yamanaka M, Brubaker S, Cawley S, Drenkow J, Piccolboni A, Bekiranov S, Helt G, Tammana H, Gingeras TR. 2004. Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res* 14:331–342. <http://dx.doi.org/10.1101/gr.2094104>.
63. Lim S, Lee B, Kim M, Kim D, Yoon H, Yong K, Kang DH, Ryu S. 2012. Analysis of HilC/D-dependent *invF* promoter expression under different culture conditions. *Microb Pathog* 52:359–366. <http://dx.doi.org/10.1016/j.micpath.2012.03.006>.
64. Alon U. 2007. Network motifs: theory and experimental approaches. *Nat Rev Genet* 8:450–461. <http://dx.doi.org/10.1038/nrg2102>.
65. Chan K, Kim CC, Falkow S. 2005. Microarray-based detection of *Salmonella enterica* serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. *Infect Immun* 73:5438–5449. <http://dx.doi.org/10.1128/IAI.73.9.5438-5449.2005>.
66. Lawley TD, Chan K, Thompson LJ, Kim CC, Govoni GR, Monack DM. 2006. Genome-wide screen for *Salmonella* genes required for long-term systemic infection of the mouse. *PLoS Pathog* 2:e11. <http://dx.doi.org/10.1371/journal.ppat.0020011>.
67. Silva-Valenzuela CA, Molina-Quiroz RC, Desai P, Valenzuela C, Porwollik S, Zhao M, Hoffman RM, Andrews-Polymenis H, Contreras I, Santiviago CA, McClelland M. 2015. Analysis of two complementary single-Gene deletion mutant libraries of *Salmonella* Typhimurium in intraperitoneal infection of BALB/c mice. *Front Microbiol* 6:1455. <http://dx.doi.org/10.3389/fmicb.2015.01455>.
68. Chaudhuri RR, Morgan E, Peters SE, Pleasance SJ, Hudson DL, Davies HM, Wang J, van Diemen PM, Buckley AM, Bowen AJ, Pullinger GD, Turner DJ, Langridge GC, Turner AK, Parkhill J, Charles IG, Maskell DJ, Stevens MP. 2013. Comprehensive assignment of roles for *Salmonella* Typhimurium genes in intestinal colonization of food-producing animals. *PLoS Genet* 9:e1003456. <http://dx.doi.org/10.1371/journal.pgen.1003456>.
69. Frye J, Karlinsey JE, Felise HR, Marzolf B, Dowidar N, McClelland M, Hughes KT. 2006. Identification of new flagellar genes of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 188:2233–2243. <http://dx.doi.org/10.1128/JB.188.6.2233-2243.2006>.
70. Hensel M. 2000. *Salmonella* pathogenicity island 2. *Mol Microbiol* 36:1015–1023. <http://dx.doi.org/10.1046/j.1365-2958.2000.01935.x>.
71. Browne SH, Lesnick ML, Guiney DG. 2002. Genetic requirements for

- Salmonella*-induced cytopathology in human monocyte-derived macrophages. *Infect Immun* 70:7126–7135. <http://dx.doi.org/10.1128/IAI.70.12.7126-7135.2002>.
72. Chakravorty D, Rohde M, Jäger L, Deiwick J, Hensel M. 2005. Formation of a novel surface structure encoded by *Salmonella* Pathogenicity Island 2. *EMBO J* 24:2043–2052. <http://dx.doi.org/10.1038/sj.emboj.7600676>.
  73. Hansen-Wester I, Stecher B, Hensel M. 2002. Type III secretion of *Salmonella enterica* serovar Typhimurium translocated effectors and SseFG. *Infect Immun* 70:1403–1409. <http://dx.doi.org/10.1128/IAI.70.3.1403-1409.2002>.
  74. Tomljenovic-Berube AM, Mulder DT, Whiteside MD, Brinkman FS, Coombes BK. 2010. Identification of the regulatory logic controlling *Salmonella* pathoadaptation by the SsrA-SsrB two-component system. *PLoS Genet* 6:e1000875. <http://dx.doi.org/10.1371/journal.pgen.1000875>.
  75. Bustamante VH, Martínez LC, Santana FJ, Knodler LA, Steele-Mortimer O, Puente JL. 2008. HilD-mediated transcriptional cross-talk between SPI-1 and SPI-2. *Proc Natl Acad Sci U S A* 105:14591–14596. <http://dx.doi.org/10.1073/pnas.0801205105>.
  76. Skinner ME, Uzilov AV, Stein LD, Mungall CJ, Holmes IH. 2009. JBrowse: a next-generation genome browser. *Genome Res* 19:1630–1638. <http://dx.doi.org/10.1101/gr.094607.109>.
  77. Anzaldi LJ, Muñoz-Fernández D, Erill I. 2012. BioWord: a sequence manipulation suite for Microsoft Word. *BMC Bioinformatics* 13:124. <http://dx.doi.org/10.1186/1471-2105-13-124>.
  78. Jarvik T, Smillie C, Groisman EA, Ochman H. 2010. Short-term signatures of evolutionary change in the *Salmonella enterica* serovar Typhimurium 14028 genome. *J Bacteriol* 192:560–567. <http://dx.doi.org/10.1128/JB.01233-09>.
  79. Stringer AM, Singh N, Yermakova A, Petrone BL, Amarasinghe JJ, Reyes-Diaz L, Mantis NJ, Wade JT. 2012. FRUIT, a scar-free system for targeted chromosomal mutagenesis, epitope tagging, and promoter replacement in *Escherichia coli* and *Salmonella enterica*. *PLoS One* 7:e44841. <http://dx.doi.org/10.1371/journal.pone.0044841>.
  80. Guzman L-M, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P<sub>BAD</sub> promoter. *J Bacteriol* 177:4121–4130.
  81. Fitzgerald DM, Bonocora RP, Wade JT. 2014. Comprehensive mapping of the *Escherichia coli* flagellar regulatory network. *PLoS Genet* 10:e1004649. <http://dx.doi.org/10.1371/journal.pgen.1004649>.
  82. Nonaka G, Blankschien M, Herman C, Gross CA, Rhodius VA. 2006. Regulon and promoter analysis of the *E. coli* heat shock factor, Sigma 32, reveals a multifaceted cellular response to heat stress. *Genes Dev* 20:1776–1789. <http://dx.doi.org/10.1101/gad.1428206>.
  83. Stringer AM, Currenti S, Bonocora RP, Baranowski C, Petrone BL, Palumbo MJ, Reilly AA, Zhang Z, Erill I, Wade JT. 2014. Genome-scale analyses of *Escherichia coli* and *Salmonella enterica* AraC reveal non-canonical targets and an expanded core regulon. *J Bacteriol* 196:660–671. <http://dx.doi.org/10.1128/JB.01007-13>.
  84. McClure R, Balasubramanian D, Sun Y, Bobrovskyy M, Sumbly P, Genco CA, Vanderpool CK, Tjaden B. 2013. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* 41:e140. <http://dx.doi.org/10.1093/nar/gkt444>.
  85. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <http://dx.doi.org/10.1101/gr.1239303>.
  86. Saldanha AJ. 2004. Java TreeView—extensible visualization of microarray data. *Bioinformatics* 20:3246–3248. <http://dx.doi.org/10.1093/bioinformatics/bth349>.
  87. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and psi-blast: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <http://dx.doi.org/10.1093/nar/25.17.3389>.
  88. Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95:14863–14868. <http://dx.doi.org/10.1073/pnas.95.25.14863>.
  89. De Hoon MJ, Imoto S, Nolan J, Miyano S. 2004. Open source clustering software. *Bioinformatics* 20:1453–1454. <http://dx.doi.org/10.1093/bioinformatics/bth078>.
  90. Grote A, Klein J, Retter I, Haddad I, Behling S, Bunk B, Biegler I, Yarmolinetz S, Jahn D, Münch R. 2009. PRODORIC (release 2009): a database and tool platform for the analysis of gene regulation in prokaryotes. *Nucleic Acids Res* 37:D61–D65. <http://dx.doi.org/10.1093/nar/gkn837>.