

# UC Irvine

## UC Irvine Previously Published Works

### Title

Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the  $\sigma^{54}$  (RpoN) regulon of Salmonella Typhimurium LT2

### Permalink

<https://escholarship.org/uc/item/63m704qb>

### Journal

BMC Genomics, 14(1)

### ISSN

1471-2164

### Authors

Samuels, David J  
Frye, Jonathan G  
Porwollik, Steffen  
[et al.](#)

### Publication Date

2013

### DOI

10.1186/1471-2164-14-602

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

RESEARCH ARTICLE

Open Access

# Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the $\sigma^{54}$ (RpoN) regulon of *Salmonella* Typhimurium LT2

David J Samuels<sup>1</sup>, Jonathan G Frye<sup>2</sup>, Steffen Porwollik<sup>3</sup>, Michael McClelland<sup>3</sup>, Jan Mrázek<sup>1,4</sup>, Timothy R Hoover<sup>1</sup> and Anna C Karls<sup>1\*</sup>

## Abstract

**Background:** Sigma54, or RpoN, is an alternative  $\sigma$  factor found widely in eubacteria. A significant complication in analysis of the global  $\sigma^{54}$  regulon in a bacterium is that the  $\sigma^{54}$  RNA polymerase holoenzyme requires interaction with an active bacterial enhancer-binding protein (bEBP) to initiate transcription at a  $\sigma^{54}$ -dependent promoter. Many bacteria possess multiple bEBPs, which are activated by diverse environmental stimuli. In this work, we assess the ability of a promiscuous, constitutively-active bEBP—the AAA+ ATPase domain of DctD from *Sinorhizobium meliloti*—to activate transcription from all  $\sigma^{54}$ -dependent promoters for the characterization of the  $\sigma^{54}$  regulon of *Salmonella* Typhimurium LT2.

**Results:** The AAA+ ATPase domain of DctD was able to drive transcription from nearly all previously characterized or predicted  $\sigma^{54}$ -dependent promoters in *Salmonella* under a single condition. These promoters are controlled by a variety of native activators and, under the condition tested, are not transcribed in the absence of the DctD AAA+ ATPase domain. We also identified a novel  $\sigma^{54}$ -dependent promoter upstream of STM2939, a homolog of the *cas1* component of a CRISPR system. ChIP-chip analysis revealed at least 70  $\sigma^{54}$  binding sites in the chromosome, of which 58% are located within coding sequences. Promoter-*lacZ* fusions with selected intragenic  $\sigma^{54}$  binding sites suggest that many of these sites are capable of functioning as  $\sigma^{54}$ -dependent promoters.

**Conclusion:** Since the DctD AAA+ ATPase domain proved effective in activating transcription from the diverse  $\sigma^{54}$ -dependent promoters of the *S. Typhimurium* LT2  $\sigma^{54}$  regulon under a single growth condition, this approach is likely to be valuable for examining  $\sigma^{54}$  regulons in other bacterial species. The *S. Typhimurium*  $\sigma^{54}$  regulon included a high number of intragenic  $\sigma^{54}$  binding sites/promoters, suggesting that  $\sigma^{54}$  may have multiple regulatory roles beyond the initiation of transcription at the start of an operon.

**Keywords:** Sigma54, RpoN, Bacterial enhancer-binding protein, Regulon, Sigma factor, *Salmonella*

\* Correspondence: akarls@uga.edu

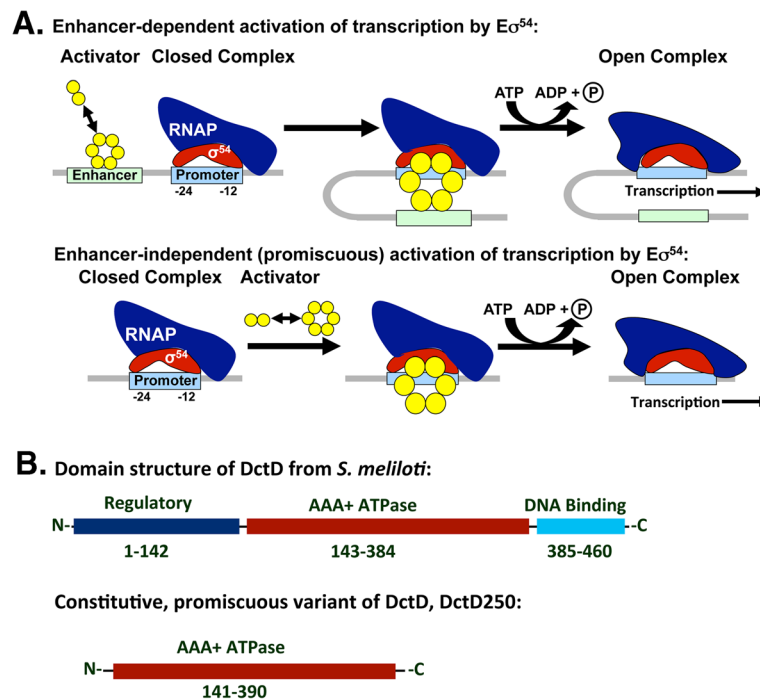
<sup>1</sup>Department of Microbiology, University of Georgia, 30602, Athens, GA, USA

Full list of author information is available at the end of the article

## Background

Transcription in eubacteria is mediated by the RNA polymerase holoenzyme ( $E\sigma$ ), which has five constant subunits ( $\alpha_2\beta\beta'\omega$ ) and a variable subunit ( $\sigma$ ). The constant subunits constitute the RNA polymerase core (RNAP), which has the polymerization activity; the  $\sigma$  subunit determines promoter recognition and functions in the  $E\sigma$ -promoter transition from closed complex to open complex (isomerization). The primary  $\sigma$  factor in a bacterium, such as  $\sigma^{70}$  in *Escherichia coli*, controls transcription of most housekeeping genes in the cell; alternative sigma factors have specialized regulons that function in the response to environmental stressors or morphological changes, or in developmental systems (for review see [1]). In many bacteria the alternative  $\sigma$  factor  $\sigma^{54}$  (also called RpoN or NtrA) has unusually diverse regulons, with genes that function in a variety of cellular processes, including flagellar biogenesis, response to nitrogen starvation, transport and metabolism of carbon substrates, and tolerance to heavy metals [2-6].

Multiple features, including protein structure, promoter consensus sequence, and mode of activation, distinguish  $\sigma^{54}$  from all other primary and secondary sigma factors, which constitute the  $\sigma^{70}$  family (reviewed in [1,7]). Although both  $\sigma^{54}$ - and  $\sigma^{70}$ -type sigma factors associate with the  $\beta$  and  $\beta'$  subunits of RNAP and mediate the binding of  $E\sigma$  to specific promoter sequences,  $\sigma^{54}$  differs extensively from  $\sigma^{70}$ -type sigma factors in primary amino acid sequence and domain organization (reviewed in [8]). The essential promoter features for  $E\sigma^{54}$  recognition and binding center around conserved GG and TGC elements at -24 and -12, respectively, relative to the transcription start site (TSS) [9], while holoenzymes with the various  $\sigma^{70}$ -type sigma factors generally recognize and bind promoter elements at -35 and -10 with the consensus sequences TTGACA and TATAAT, respectively (reviewed in [1]). Perhaps the most important feature of  $E\sigma^{54}$  that differs from  $E\sigma^{70}$  is the isomerization process (Figure 1A). For  $E\sigma^{70}$  the transition from closed complex to open complex is usually spontaneous and rapid, so regulation of trans-



**Figure 1** Activation of  $\sigma^{54}$ -dependent transcription and activator structure. **A)**  $\sigma^{54}$  (red subunit) directs binding of the RNA polymerase (dark blue subunit) holoenzyme ( $E\sigma^{54}$ ) to the -12, -24 promoter elements (light blue box). This closed complex is stable and cannot transition to open complex. In response to an environmental or cellular signal, the activator (bEBP; yellow dimers) oligomerizes. For most bEBPs, the oligomer binds to an enhancer (green box) 80 to 150 bp upstream of the promoter and DNA looping brings the activator in contact with  $\sigma^{54}$  in the  $E\sigma^{54}$  closed complex. Hydrolysis of ATP by bEBP causes remodeling of  $E\sigma^{54}$ , which leads to open complex formation and transcription. There are a few bacteria with bEBPs that are missing the DNA binding domain; after oligomerization, these activators can bind to  $\sigma^{54}$  in closed complex with any promoter to stimulate open complex formation (promiscuous activation). **B)** The domain structure for the *Sinorhizobium meliloti* bEBP, DctD, is typical of most bEBPs. The amino-terminal regulatory domain (dark blue box) inhibits assembly of the bEBP oligomer until it interacts with an activation signal; the AAA+ ATPase domain (red box) mediates ATP binding and hydrolysis, as well as the protein-protein interactions between bEBPs (oligomerization) and between bEBP and  $\sigma^{54}$ ; the carboxyl-terminal DNA binding domain (aqua box) contains a helix-turn-helix motif for binding the enhancer. The truncated DctD variant, DctD250, is missing the regulatory and DNA binding domains, so that it is constitutively active and promiscuous in stimulating transcription from  $\sigma^{54}$ -dependent promoters.

cription initiation frequently occurs at the level of closed complex formation. Initiation of transcription by  $E\sigma^{54}$  more closely resembles eukaryotic Pol II systems in that  $E\sigma^{54}$  forms a stable closed complex that requires a bacterial enhancer-binding protein (bEBP) and ATP hydrolysis for isomerization to open complex (reviewed in [10]). The bEBPs add an additional level of complexity to the  $\sigma^{54}$  regulon.

bEBPs have a modular structure that is generally conserved: an N-terminal regulatory domain, a central AAA+ ATPase domain, and a C-terminal DNA binding domain (Figure 1B; reviewed in [8]). These proteins activate transcription from  $\sigma^{54}$ -dependent promoters in three basic steps (Figure 1A). First, the bEBP receives an environmental stimulus through phosphorylation, ligand binding, or protein-protein interactions with the N-terminal regulatory domain that stimulates the bEBP to multimerize through the AAA+ ATPase domain and bind to an upstream activator sequence (UAS or enhancer) via the C-terminal DNA binding domain. The bEBP-UAS complex is then brought into contact with the  $E\sigma^{54}$ -promoter closed complex via a DNA looping event and interactions between highly conserved regions of the AAA+ ATPase domain of bEBP and  $\sigma^{54}$ . Finally, ATP hydrolysis drives isomerization, allowing the initiation of transcription.

The requirement for bEBP-mediated activation of  $\sigma^{54}$ -dependent transcription presents two problems for global analysis of a  $\sigma^{54}$  regulon. The first is the need for the proper environmental stimulus to activate bEBPs. Since the  $E\sigma^{54}$  closed complex requires an activated bEBP,  $\sigma^{54}$ -dependent promoters are usually transcriptionally silent in the absence of the specific stimulus for the bEBP [8]. Analysis of transcription from  $\sigma^{54}$ -dependent promoters under any single growth condition would miss operons whose bEBPs are not activated under the condition tested. Secondly, the requirement for the UAS or enhancer by most bEBPs presents a challenge for predicting whether a  $E\sigma^{54}$  binding site is functioning as a promoter or not. There is no common consensus sequence for the enhancer and their position relative to the promoter can be quite variable. For many  $\sigma^{54}$ -dependent promoters the UAS sequence lies ~70-150 bp upstream of the promoter, but other configurations have been characterized, such as enhancers located 1.5 kb downstream of the *rocG* promoter in *Bacillus subtilis* [11] and up to 3 kb upstream of the promoter in artificial constructs of the *glnA* operon from *E. coli* [12]. If a  $\sigma^{54}$  binding site is examined for promoter activity in isolation, such as in a promoter-reporter vector, it is difficult to discern whether a site is inactive because it is not a promoter or because the enhancer was not included in the cloned sequence.

Previous studies to define the  $\sigma^{54}$  regulons of *Escherichia coli* [13], *Vibrio cholerae* [14] and *Geobacter sulfurreducens* [15] have recognized the limitations presented by the

requirement for activated bEBPs in the characterization of the full  $\sigma^{54}$  regulon, even when  $\sigma^{54}$  is overexpressed from a heterologous promoter. Our approach to overcoming these problems in the global characterization of  $\sigma^{54}$  regulons in bacteria is the utilization of a constitutively-active, promiscuous bEBP, the AAA+ ATPase domain of *Sinorhizobium meliloti* DctD [16,17]. We chose to assess the efficacy of this approach in *Salmonella enterica subsp. enterica* serovar Typhimurium LT2 (hereafter referred to as *S. Typhimurium* LT2), a widely-used laboratory strain, because it has a moderately-sized  $\sigma^{54}$  regulon with 13 known or predicted bEBPs [18], providing sufficient diversity in bEBPs to test our hypothesis.

We report here that use of this constitutively-active, promiscuous bEBP in DNA microarrays and promoter function assays permitted detection of nearly all known and predicted  $\sigma^{54}$ -dependent operons. These studies also revealed a new  $\sigma^{54}$ -dependent promoter expressing a putative *cas1* gene in *S. Typhimurium* LT2 (STM2938). In addition, chromatin immunoprecipitation-microarray (ChIP-chip) analysis combined with bioinformatics identified 70  $E\sigma^{54}$  or  $\sigma^{54}$  binding sites, of which 41 appear to be within open reading frames (ORFs). This surprising number of intragenic sites suggests regulatory roles for  $\sigma^{54}$  or  $E\sigma^{54}$  that may involve repression, transcriptional interference, or expression of cis- or trans-acting small RNA (sRNA) [19,20].

## Results and discussion

### Utility of a promiscuous, constitutive bEBP in characterizing the $\sigma^{54}$ regulon

Since all known  $\sigma^{54}$ -dependent promoters require an activated bEBP for transcription initiation, it is a challenge to find a condition under which all promoters can be detected within the  $\sigma^{54}$  regulon of a bacterium. In the recent mapping of the *S. Typhimurium* SL1344 transcriptome using early stationary phase cultures in rich media (Lennox broth), only one of the known or predicted  $\sigma^{54}$ -dependent gene transcripts was detected, *pspA* [21]. The currently favored approach is overexpression of  $\sigma^{54}$  to facilitate detection of  $\sigma^{54}$ -dependent promoters, which assumes a reasonable basal level of activation of the bEBPs. Using relatively low cutoffs for the fold-change (1.5- to 2-fold) in transcript levels between the  $\sigma^{54}$ -overexpression strain and wild type or  $\Delta rpoN$  strains, a considerable portion of the  $\sigma^{54}$ -dependent transcriptome was defined in *Escherichia coli* [13], *Vibrio cholerae* [14] and *Geobacter sulfurreducens* [15]. However, not all previously-identified  $\sigma^{54}$ -dependent operons were detected for *E. coli* and *G. sulfurreducens*, and evidence from the *V. cholerae* and *G. sulfurreducens* studies suggests that overexpression of  $\sigma^{54}$  may repress expression from some  $\sigma^{54}$ -dependent promoters and alter expression of  $\sigma^{54}$ -independent promoters [13-15]. We hypothesize that

a promiscuous and constitutive variant of the bEBP DctD from *S. meliloti* can activate transcription from all  $\sigma^{54}$ -dependent promoters in *S. Typhimurium* LT2 at wild-type levels of  $\sigma^{54}$  under a single growth condition, thereby facilitating global characterization of the  $\sigma^{54}$  regulon without overexpression of  $\sigma^{54}$ . This promiscuous and constitutive DctD variant is missing the N-terminal response regulator and C-terminal DNA binding domains, leaving only the central AAA+ ATPase domain, residues 141 to 390 of DctD and referred to hereafter as DctD250 [17]. Previous work showed that DctD250 was able to interact with  $E\sigma^{54}$  in *E. coli* to drive transcription from the chromosomal *glnA* promoter and from the *S. meliloti* *dctA* promoter in the absence of native DctD and without an enhancer sequence [16,17].

The  $\sigma^{54}$ -dependent promoters of *S. Typhimurium* LT2 are normally responsive to one or more of thirteen known and predicted bEBPs under various growth conditions [18], so to initially assess DctD250 activation of transcription from  $\sigma^{54}$ -dependent promoters that respond to different bEBPs in *Salmonella*, the  $\sigma^{54}$ -dependent promoters for the *glnKamtB* (STM0462) and *rtcBA* (STM3521) operons were introduced upstream of a promoter-less *lacZ* gene and the reporter plasmids were transformed into a derivative of *S. Typhimurium* LT2 (wild-type; WT) and WT containing the DctD250 expression plasmid (WT + DctD250) to perform  $\beta$ -galactosidase assays. The *glnKamtB* and *rtcBA* promoters were chosen because neither has predicted  $\sigma^{70}$ -dependent promoters within the cloned promoter region and each is responsive to a different bEBP: NtrC for *glnKamtB* [22] and RtcR for *rtcBA* [23]. In the WT strain, the *glnKamtB* and *rtcBA* operon promoters expressed *lacZ* at very low levels; but in the presence of DctD250, *lacZ* was expressed at 150- and 16-fold higher levels, respectively (Table 1). To compare the level of expression stimulated by DctD250 to the level that is seen under physiological conditions that activate the promoter-associated bEBP, *lacZ* expression from the *glnKamtB* promoter was assayed in the WT strain in nitrogen-limiting medium, which activates NtrC. Under nitrogen-starvation conditions NtrC multimerizes, binds the enhancer in the cloned promoter region, and hydrolyzes ATP to stimulate transcription by  $E\sigma^{54}$  at the *glnKamtB* promoter (see Figure 1A). In the presence of activated NtrC, the *glnKamtB* promoter expresses *lacZ* at a nearly 10-fold higher level than in the presence of DctD250. This reduced level of activation by DctD250 relative to the cognate bEBP under activation conditions is consistent with previous studies comparing the activity of truncated versions of bEBPs, which must interact with  $E\sigma^{54}$  from solution, to that of the wild type bEBPs, which are directed to the target  $\sigma^{54}$  promoter via binding to the enhancer sequence [17,24]. The control reporter plasmids pDV6, which has the  $\sigma^{70}$ -dependent, circle junction promoter

**Table 1 DctD250-dependent activity of predicted and potential  $\sigma^{54}$ -dependent promoters**

Promoter <sup>a</sup>	Miller units <sup>b</sup>		
	WT	WT+DctD250	WT-N
<b>Intergenic:</b>			
STM0462 ( <i>glnKamtB</i> )	1.1 ±0.1	180 ±22	1778 ±28
STM3521 ( <i>rtcBA</i> )	14 ±6.8	220 ±55	N.D.
<b>Intragenic:</b>			
STM0699	5.3 ±1.6	56 ±6.2	N.D.
STM2430	19 ±3.8	90 ±5.5	N.D.
STM2939	2.2 ±1.4	76 ±17	N.D.
<b>Controls:</b>			
IS492-CJ (pDV6)	6400 ±2200	9000 ±4100	N.D.
Empty vector (pDS12)	2.0 ±0.5	8.4 ±3.5	4 ±0

<sup>a</sup>Promoters assayed in *lacZ*-reporter plasmids pDS11 and pDS12: intergenic promoters are predicted promoters for the *glnALG*, *glnKamtB* and *rtcBA* operons; intragenic promoters were identified in the ChIP-chip assay (see Results). Controls are the  $\sigma^{70}$ -dependent, circle-junction promoter from IS492 and the empty vector, pDS12.

<sup>b</sup> $\beta$ -galactosidase assays were performed in MOPS minimal medium (WT and WT+DctD250) or in nitrogen-limiting MOPS (WT-N).

from IS492 [25], and the promoter-less pDS12 expressed *lacZ* at approximately the same level in WT as WT + DctD250 (Table 1). Based on these results, DctD250 activates transcription from  $\sigma^{54}$ -dependent promoters that are normally responsive to different bEBPs under different growth conditions. Therefore, we performed DNA microarray and promoter-reporter analyses in the presence of the promiscuous, constitutive activator DctD250 to assess the efficacy of this approach in defining the  $\sigma^{54}$  regulon of *S. Typhimurium* LT2.

#### Microarray analysis of $\sigma^{54}$ -dependent transcripts in *Salmonella* expressing DctD250

To determine the genes whose transcription is controlled by  $\sigma^{54}$  in *S. Typhimurium* LT2 we performed a microarray analysis comparing WT+DctD250 to an isogenic strain with a deletion of *rpoN* ( $\Delta rpoN$ +DctD250). RNA collected during mid-log phase growth in nutrient medium was reverse transcribed and cDNAs from each strain were differentially labeled and applied to a complete ORF array containing all annotated open reading frames for *S. Typhimurium* LT2 [26]. Open reading frames that were transcribed in WT at a level > 3-fold higher than in the  $\Delta rpoN$  strain, with a p-value <0.02, were considered up-regulated and, for the purpose of the initial categorization of these results, an operon was considered up-regulated if at least one gene met these criteria. In three biological replicates, the same 33 operons were up-regulated in the presence of  $\sigma^{54}$ . The microarray results for *S. Typhimurium* LT2 genes within operons that meet the criteria for up-regulation, or that are known or predicted to be  $\sigma^{54}$ -dependent, are shown in Table 2 and Additional file 1.

**Table 2 Microarray results for known, predicted, and novel  $\sigma^{54}$ -dependent operons and sRNA genes of *S. Typhimurium***

Locus tag <sup>a</sup>	Gene symbol <sup>b</sup>	Function	bEBP <sup>c</sup>	WT/ $\Delta$ rpoN <sup>d</sup>	Ref. <sup>e</sup>
<b>Known <math>\sigma^{54}</math>-dependent operons and sRNA genes:</b>					
<b>STM0368-71</b>	<i>prpBCDE</i>	Propionate catabolism (putative)	PrpR	<b>45</b>	[4]
<b>STM0830-28</b>	<i>glnHPQ</i>	Glutamine high-affinity transporter	NtrC	<b>7.1</b>	[27]
<b>STM2355</b>	<i>argT</i>	Lysine/arginine/ornithine transport protein	NtrC	<b>3.5</b>	[2]
STM_R0152	<i>glmY</i>	GlmY sRNA	GlrR	0.9	[28]
STM_R0167	<i>glmZ</i>	GlmZ sRNA	GlrR	1.1	[28]
<b>STM4007-05</b>	<i>glnALG</i>	Glutamine synthetase	NtrC	<b>48</b>	[29]
<b>Predicted <math>\sigma^{54}</math>-dependent operons:</b>					
<b>STM0462-63</b>	<i>glnK amtB</i>	hypothetical protein	NtrC	<b>3.6</b>	[22]
<b>STM0577-72</b>		PTS (putative)	STM0571	<b>67</b>	[18]
<b>STM0649.S-53</b>		Hydrolase (putative)	STM0652	<b>11</b>	[18]
STM0665-62	<i>gltJJKL</i>	Glutamate/aspartate transporter	NtrC	1.8	[18,30]
STM1285-84	<i>yeaGH</i>	Serine protein kinase (putative)	NtrC	2.5	[18,30]
<b>STM1303-07</b>	<i>astCABDE</i>	Arginine/ornithine/glutamine metabolism	NtrC	<b>2.4<sup>f</sup></b>	[31,32]
<b>STM1690-86</b>	<i>pspABCDE</i>	Phage shock proteins	PspF	<b>17</b>	[5]
<b>STM2360-56</b>	----- <i>ubiX</i>	Amino acid transport (putative)	STM2361	<b>100</b>	[18]
<b>STM2840-41</b>	<i>norV ygbD</i>	Nitric oxide reductase	NorR	<b>16</b>	[18,33]
<b>STM2843-42</b>	<i>hydN hypF</i>	Hydrogenase maturation proteins	FhIA	<b>13</b>	[34]
<b>STM2853-44</b>	<i>hycABCDEFGH-</i>	Hydrogenase 3	FhIA	<b>26</b>	[35]
<b>STM2854-58</b>	<i>hypABCDE</i>	Formate-hydrogen lyase system	FhIA	<b>5.6</b>	[35]
<b>STM3521-18</b>	<i>-rtcBA</i>	RNA repair system (putative)	RtcR	<b>71</b>	[23]
STM3568	<i>rpoH</i>	Heat shock sigma factor ( $\sigma^{32}$ )		1.7	[36,37]
<b>STM3772-66</b>		PTS (putative)	STM3773	<b>39</b>	[18]
<b>STM4172</b>	<i>zraP</i>	Zinc resistance-associated protein	ZraR	<b>16</b>	[3,18]
<b>STM4173-74</b>	<i>hydHG</i>	Zinc resistance two-component system	ZraR	<b>3.7</b>	[3]
STM4244	<i>pspG</i>	Phage shock protein	PspF	1.4	[38]
<b>STM4285</b>	<i>fdhF</i>	Formate dehydrogenase	FhIA	<b>29</b>	[39]
<b>STM4535-40.s</b>		PTS (putative)	STM4534	<b>16</b>	[18]
<b>Novel <math>\sigma^{54}</math>-dependent operon:</b>					
<b>STM2944-2937</b>		CRISPR-associated genes		<b>1.6<sup>f</sup></b>	-

<sup>a</sup>Locus tags for genes within operons or sRNA genes are grouped by those previously shown to be  $\sigma^{54}$ -dependent in *Salmonella*, previously predicted to be  $\sigma^{54}$ -dependent, or identified in this study as encoded in a novel  $\sigma^{54}$ -dependent transcript. Locus tags for operons that are not up-regulated are in bold type.

<sup>b</sup>Genes that have not been assigned a gene symbol are represented by a dash (-).

<sup>c</sup>Known or predicted bacterial enhancer-binding protein (bEBP) that activates the  $\sigma^{54}$ -dependent operon.

<sup>d</sup>Signal ratio for the first gene in the operon in WT and  $\Delta$ rpoN strains expressing DctD250 from pPHBP92. Operons with at least one gene with a signal ratio >3 and p-value <0.02 are considered up-regulated by RpoN; signal ratios above the 3-fold cut off are in bold type. Data for all genes in these operons can be found in Additional file 1.

<sup>e</sup>References for operons shown to be  $\sigma^{54}$ -dependent in *Salmonella* and for operons either determined to be  $\sigma^{54}$ -dependent in other bacterial genera or predicted to be regulated by  $\sigma^{54}$  in *Salmonella* are listed.

<sup>f</sup>The first gene in the operon was <3-fold up-regulated, but other genes in the operon were >3-fold up-regulated.

Only 4 genes, STM2722, STM2724, STM2729, and STM2730, which are part of 2 operons in the Fels-2 prophage, were down-regulated >3-fold with a p-value <0.02 in the WT strain as compared to the  $\Delta$ rpoN strain.

#### Known $\sigma^{54}$ -dependent operons and sRNA

If our hypothesis is correct, then in the presence of DctD250 we should observe up-regulation of operons (one or more structural genes) and sRNA genes that are

known to have  $\sigma^{54}$ -dependent promoters, even though they are normally activated by different bEBPs. Previously, four *Salmonella* operons have been experimentally shown to be regulated by  $\sigma^{54}$ : *prpBCDE* [4], *glnHPQ* [27], *argT* [2], and *glnALG* [29]. Additionally, two sRNA genes, *glmY* and *glmZ*, have also been shown to have  $\sigma^{54}$ -dependent promoters [28]. Table 2 summarizes the genes, functions, bEBPs, and microarray results for the known  $\sigma^{54}$ -dependent operons and sRNA genes of *Salmonella*.

The DNA microarrays showed up-regulation of all four known  $\sigma^{54}$ -dependent operons in *Salmonella*, *prpBCDE*, *glnHPQ*, *argT*, and *glnALG* (Table 2). The two sRNA genes with known  $\sigma^{54}$ -dependent promoters did not appear up-regulated by  $\sigma^{54}$ . This result was not surprising since in *S. Typhimurium* both *glmY* and *glmZ* possess  $\sigma^{70}$ -dependent promoters that fully overlap the  $\sigma^{54}$ -dependent promoters, such that the  $E\sigma^{70}$  and  $E\sigma^{54}$  compete for binding to their respective promoters [28]. Gopel et al. [28] demonstrated that the level of *glmY* transcription was similar in wild type and  $\Delta rpoN$  cells and that transcription of *glmZ* actually increased in the *rpoN* mutant, reflecting that the  $\sigma^{70}$ -dependent promoter for *glmZ* is stronger than the  $\sigma^{70}$ -dependent promoter for *glmY*. The presence of a  $\sigma^{70}$  promoter does not necessarily preclude detection of a  $\sigma^{54}$ -dependent promoter controlling expression of a gene or operon in these microarray assays, though; the promoter region of *glnA* has non-overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -dependent promoters [29], yet was up-regulated 48-fold. Taken together, these results for the known  $\sigma^{54}$ -dependent promoters are consistent with our hypothesis that DctD250 can promiscuously and constitutively activate  $\sigma^{54}$ -holoenzyme at a variety of  $\sigma^{54}$ -dependent promoters.

#### Confirmation of predicted $\sigma^{54}$ -dependent operons

There are 20 operons that we define as 'predicted'  $\sigma^{54}$ -dependent operons in *Salmonella*. These predictions are based on *in silico* analyses indicating either homology to known  $\sigma^{54}$ -dependent operons in *E. coli* and other enteric bacteria or promoter sequence homology along with genetic proximity to predicted bEBP genes [3,5,18,22,23,30-39]. However,  $\sigma^{54}$ -dependent transcription of these operons has not previously been experimentally demonstrated in *Salmonella*. In the DNA microarrays, 16 of the 20 operons that have been predicted to have  $\sigma^{54}$ -dependent promoters in *Salmonella* were up-regulated in WT+DctD250 as compared to  $\Delta rpoN$ +DctD250 (Table 2), providing experimental evidence that these genes are, in fact, regulated by  $\sigma^{54}$  in *S. Typhimurium* LT2.

For these 16 up-regulated  $\sigma^{54}$ -dependent operons there are 11 different bEBPs that either are known or predicted to activate expression from their  $\sigma^{54}$ -dependent promoters (Table 2). Five of the up-regulated operons, STM0577-0572, STM0649.s-0653, STM2360-2356, STM3772-3766, and STM4535-4540.s, were predicted to be  $\sigma^{54}$ -dependent based on linkage to a predicted bEBP and an upstream sequence with the essential -12 and -24 elements of a  $\sigma^{54}$ -dependent promoter [18]. There are no orthologs in *E. coli* for the predicted bEBPs associated with these operons; three of these predicted bEBPs, STM0571, STM3773 and STM4534, are similar to the LevR-type EBPs found in Gram-positive bacteria [18]. In

addition to the microarray evidence presented here for  $\sigma^{54}$  regulation of these operons, we know that STM3773 is the bEBP controlling expression of STM3772-3776 and that this operon encodes the components of a phosphotransferase system permease for D-glucosaminic acid and enzymes required for catabolism of this acid sugar [40]. These results show that DctD250 can activate expression at  $\sigma^{54}$ -dependent promoters that are normally regulated by the LevR-type bEBPs.

Of the four predicted  $\sigma^{54}$ -dependent operons that did not fulfill our criteria for upregulation in the microarray, at least two have additional  $\sigma^{54}$ -independent promoters, which may have masked the effect of  $\sigma^{54}$  on transcription levels. The heat shock sigma factor gene *rpoH* has been shown to be under the control of additional promoters and other regulatory proteins in *E. coli* [36]. The conservation of this promoter region for *rpoH* in *S. Typhimurium* LT2 suggests that a similar complex regulatory scheme may be involved [37], thereby reducing the effects of the  $\Delta rpoN$  mutation. The *yeaGH* operon, which was just below the 3-fold cutoff for up-regulation in the microarray analysis, has previously been shown to be under control of  $\sigma^S$  in *Salmonella* [41]; however, our assays utilized *S. Typhimurium* LT2, which has a defective *rpoS* gene due to a transversion mutation in the start codon [42]. The promoter-reporter assay with the *yeaGH* promoter region, described below, suggests there is a  $\sigma^{54}$ - and  $\sigma^S$ -independent promoter expressing the *yeaGH* operon in both the WT+DctD250 and  $\Delta rpoN$ +DctD250 strains.

The frequency of alternate promoters seen for the  $\sigma^{54}$ -dependent operons in *Salmonella* (at least 15% for the known and predicted promoters in our analyses) is not unique. Zhao et al. [13] estimate that 14% of  $\sigma^{54}$ -dependent genes in *E. coli* are transcribed by  $\sigma^{70}$ -associated RNA polymerase and suggest that expression of  $\sigma^{54}$ -dependent genes from alternate promoters allows for differential expression under various environmental conditions.

#### New potential $\sigma^{54}$ -dependent genes

In addition to the  $\sigma^{54}$ -dependent expression of known or predicted genes and operons, the DNA microarray analysis revealed up-regulation of a gene, STM2938, which has not previously been reported or predicted to be  $\sigma^{54}$ -dependent. STM2938 is the penultimate gene in a nine-gene operon that is annotated as a group of CRISPR-associated (*cas*) genes. Although none of the other genes in this operon seem to be controlled by  $\sigma^{54}$ , further evidence is presented below that supports the presence of a  $\sigma^{54}$ -dependent promoter within the gene upstream of STM2938. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and *cas* genes constitute an

adaptive immune system in bacteria and archaea that protects against invading mobile DNA, such as phage and plasmids [43]. The response to phage infection, which is referred to as phage shock, is regulated by  $\sigma^{54}$  and the bEBP PspF in *E. coli* [5]; thus, it would not be surprising for essential components of the bacterial immune response in phage infection to be regulated similarly. The potential  $\sigma^{54}$ -dependent gene STM2938 is a homologue of the *cas1* gene, which is an endonuclease that is associated with all CRISPR loci and is most likely involved in the adaptation phase of the CRISPR-immune system [44]. The regulation of this *cas1*-like gene by PspF in *Salmonella* is currently under investigation.

There were 12 additional ORFs that met the 3-fold cutoff for up-regulation by  $\sigma^{54}$  in the microarray assay, including genes for pilin biosynthesis (*hofB*), histidine ammonia lyase (*hutH*), bEBPs (*ygaA*, *fhlA*), propanediol utilization (*pduG*), siderophore production (*iroD*), and cell invasion (*invG*). The whole genome chromatin immunoprecipitation assays described below did not reveal  $\sigma^{54}$  binding sites associated with these ORFs; thus the expression of these genes may be indirectly affected by the absence of  $\sigma^{54}$  in the  $\Delta rpoN$  mutant, or constitute false positives (Additional file 1).

#### ChIP-chip analysis of genome-wide $\sigma^{54}$ binding sites in *Salmonella*

In the characterization of the  $\sigma^{54}$  regulon of *Salmonella*, determination of the genomic binding sites for the  $E\sigma^{54}$  allows confirmation of primary transcripts indicated by microarray analysis and recognition of potential  $\sigma^{54}$ -regulated genes that might not have been detected due to instability of the transcripts. To assess the binding of  $E\sigma^{54}$  in the *S. Typhimurium* LT2 genome, we isolated  $\sigma^{54}$ -DNA complexes from WT and  $\Delta rpoN$  strains that did not contain the DctD250 expression plasmid. Since bEBPs do not activate transcription by recruiting  $E\sigma^{54}$  to promoter sequences [8], inclusion of DctD250 should not be necessary to detect binding of holoenzyme to promoter sequences in the ChIP-chip assay. Protein-DNA complexes containing either  $E\sigma^{54}$  or  $\sigma^{54}$  are pulled down in the ChIP with  $\alpha\text{-}\sigma^{54}$ . The  $\sigma^{54}$  subunit is most likely to interact with the genome in the context of the RNA polymerase holoenzyme; however,  $\sigma^{54}$  has been shown to specifically bind in the absence of the core RNA polymerase at  $\sigma^{54}$ -dependent promoters that have a T-tract upstream of the GC in the -12 promoter element [45]. DNA fragments from the  $\alpha\text{-}\sigma^{54}$  ChIP were labeled and applied to the same complete open reading frame arrays as used in the microarray analysis.

Since the use of the ORF arrays did not allow direct mapping of the binding sites, we combined the ChIP-chip data with *in silico* analysis to determine the potential  $\sigma^{54}$  DNA binding sites. A Position-Specific Score

Matrix (PSSM) was created using 27 known or previously predicted  $\sigma^{54}$ -dependent promoters from *S. Typhimurium* LT2 (Additional file 2); the extent of each promoter sequence used for the PSSM (18 bp) was based on the consensus sequence for  $\sigma^{54}$ -dependent promoters defined by Barrios et al. [9] and comparison analysis of the known *Salmonella*  $\sigma^{54}$ -dependent promoters. This PSSM was applied with the Motif Locator program [46] to the enriched ORF sequence and 1000 bp of flanking sequence on both sides of the ORF to identify potential  $\sigma^{54}$  DNA binding sites. The size range of DNA fragments that were pulled down via ChIP and amplified by ligation-mediated PCR was 200–1000 bp long, as determined by agarose gel electrophoresis, suggesting that intergenic binding sites up to 1000 bp from the enriched ORF might be detected in the ChIP-chip assays.

#### $\sigma^{54}$ binding to promoters for known, predicted, and novel $\sigma^{54}$ -dependent operons

In the ChIP-chip assays with the WT and  $\Delta rpoN$  strains, the promoter-proximal gene for all the 24 known and predicted  $\sigma^{54}$ -dependent operons and the 2 sRNA genes (Table 2) were enriched, as defined by a stringent cut-off, i.e. signal ratio  $\geq 3$  and p-value  $< 0.02$  (Table 3). The associated promoter sequences, as determined by the *in silico* analysis, had PSSM scores ranging from 10.9 to 23.6 and were within 27 to 154 bp of the enriched ORF. In the DNA microarrays, six of the known or predicted  $\sigma^{54}$ -dependent operons did not appear up-regulated; but in the ChIP-chip assays the promoter regions for all six operons gave signal ratios ranging from 3.2- to 39-fold greater in WT than in  $\Delta rpoN$  cells. The detection of the  $\sigma^{54}$ -dependent promoters for all the other known and predicted  $\sigma^{54}$ -dependent operons supports the efficacy of our approach to mapping potential  $\sigma^{54}$ -binding sites.

The ChIP-chip analyses also showed that only one (STM2938) of the 13 newly-identified, potential  $\sigma^{54}$ -dependent operons from the DNA microarray assays has a  $\sigma^{54}$  DNA binding site associated with it (Table 3), suggesting that the other 12 operons may be indirectly regulated by  $\sigma^{54}$ . The  $\sigma^{54}$  DNA binding site associated with STM2938, the *cas1*-like gene, is within the upstream gene, STM2939 (539 bp from the start of STM2938). Further characterization of this potential  $\sigma^{54}$  promoter is described below in the promoter-reporter analysis.

#### $\sigma^{54}$ binding to newly identified potential promoter and regulatory sites

In total, 70 ORFs were each found to be enriched in 3 replicate samples for the WT cells as compared to the  $\Delta rpoN$  cells in the ChIP-chip assays (Table 3). The potential  $\sigma^{54}$  binding site with the highest PSSM score within each enriched ORF or up to 1000 bp of flanking



**Table 3** ChIP-chip signal ratios, PSSM scores, and predicted binding sites for ORFs enriched in the presence of  $\sigma^{54}$

Locus tag <sup>a</sup>	Gene name	Signal ratio <sup>b</sup>	Orientation <sup>c</sup>	PSSM <sup>d</sup>	Start	End	Sequence
<i>Sites located within intergenic regions:</i>							
<b>STM0368</b>	<i>prpB</i>	11	+	20.7	417914	417931	TGGCATAGCCTTTGCTTT
STM0448	<i>clpP</i>	4.6	+	14.6	503028	503045	TGTCACGTATTTTGCAATG
<b>STM0462</b>	<i>glnK</i>	3.2	+	20.0	520445	520462	TGGCACATCCTTTGCAAT
<b>STM0577</b>		8.3	+	17.9	636883	636866	TGGCACGCCGTTTGCCAT
<b>STM0649.S</b>		6.9	+	18.7	711945	711962	TGGCACGCCTTTTGATTA
<b>STM0665</b>	<i>gltI</i>	3.2	+	22.1	730107	730090	TGGCACGTCTATTGCTTT
<b>STM0830</b>	<i>glnH</i>	16	+	20.8	897079	897062	TGGCATGATTTTTTCATT
<b>STM1285</b>	<i>yeaG</i>	4.1	+	21.5	1363884	1363867	TGGCATGAGAGTTGCTTT
<b>STM1303</b>	<i>astC</i>	4.0	+	21.7	1382105	1382122	TGGCACGAATGCTGCAAT
<b>STM1690</b>	<i>pspA</i>	23	+	20.3	1782486	1782469	TGGCACGCAAATTGTATT
<b>STM2355</b>	<i>argT</i>	3.5	+	16.2	2466359	2466376	TGGCATAAGACCTGCATG
<b>STM2360</b>		4.8	+	23.6	2472731	2472714	TGGCATGCCTTTTGCTTT
<b>STM_R0152</b>	<i>glmY</i>	31	+	20.6	2707874	2707857	TGGCACAAATTAATGCATA
STM2809	<i>proV</i>	9.5	-	14.6	2955839	2955822	TGGCATGAATATTGCGAG
<b>STM2840</b>		6.5	+	20.1	2985009	2985026	TGGCACACTAGCTGCAAT
<b>STM2843</b>	<i>hydN</i>	23	+	17.1	2990721	2990704	TGGCACGATTCGTGTATA
<b>STM2853</b>	<i>hycA</i>	31	+	17.9	2999639	2999622	TGGCATGGAAAATGCTTA
<b>STM2854</b>	<i>hypA</i>	71	+	22.4	2999753	2999770	TGGCATAAATATTGCTTT
<b>STM3521</b>	<i>rsr</i>	15	+	21.2	3684734	3684717	TGGCACGCTGGTTGCAAT
<b>STM3568</b>	<i>rpoH</i>	22	+	18.9	3736836	3736819	TGGCACGGTTGTTGCTCG
<b>STM3772</b>	<i>dgaA</i>	3.6	+	20.3	3972484	3972467	TGGCACAACTTTGCTCT
<b>STM_R0167</b>	<i>glmZ</i>	15	+	19.5	4141620	4141637	TGGCACGTTATGTGCAAT
<b>STM4007</b>	<i>glnA</i>	4.2	+	19.2	4217110	4217093	TGGCACAGATTTGCTTT
<b>STM4172</b>	<i>zraP</i>	29	+	17.4	4388217	4388234	TGGCACGGAAGATGCAAG
<b>STM4173</b>	<i>hydH</i>	4.8	+	20.1	4388385	4388402	TGGCATGATCTCTGCTTA
<b>STM4244</b>	<i>pspG</i>	39	+	19.4	4465042	4465059	TGGCATGATTTTTGTAAAG
<b>STM4285</b>	<i>fdhF</i>	10	+	18.2	4527564	4527547	TGGCATAAAACATGCATA
STM4367	<i>yjeB</i>	3.8	+	14.1	4610407	4610424	TGGCAGATATTTGCTTG
<b>STM4535</b>		12	+	18.3	4794881	4794898	TGGCACGCCGCTTGTCTCT
<i>Sites located within the enriched ORF:</i>							
STM0131	<i>ftsQ</i>	7.6	+	6.2	153598	153615	TGGAACGCGTCTTGCAAG
STM0155		4.1	+	9.5	182767	182784	CGGCATGGCATTGCCAG
STM0322	<i>proA</i>	7.8	-	11.3	368058	368041	CGGCACAGTTATGCAAG
STM0332		3.0	-	8.1	376286	376269	TGGCCAGAAATATGCTTA
STM0526	<i>ylbA</i>	4.3	+	9.1	588233	588216	TGGCATTAAATGCTGCATC
STM0699		14	+	13.7	761691	761674	TGGCATCGATATTGCAAA
STM0879 <sup>^</sup>	<i>potH</i>	5.2	+	12.1	951550	951567	TGGCAGGAGTTTTTCAAT
STM0884	<i>ulaA</i>	5.2	+	10.0	955545	955562	CGGCACGATTTTTTCCAT
STM0901		3.9	+	11.7	971761	971778	TGGCATGAAACTTGTCAC
STM0940		9.5	+	13.7	1018097	1018080	TGGCCTGAATCTTGCTAA
STM0961		7.6	-	17.7	1041686	1041669	TGGCATGAAAGCTGCTCA
STM1361	<i>ydiM</i>	3.6	+	11.6	1443903	1443886	TGGCATTCTTTATGCTCA
STM1390	<i>orf242</i>	8.9	-	12.9	1475563	1475546	TGGCATCATTATTGCCTA

**Table 3** CHIP-chip signal ratios, PSSM scores, and predicted binding sites for ORFs enriched in the presence of  $\sigma^{54}$  (Continued)

STM1409	<i>ssaJ</i>	5.0	+	6.5	1490273	1490290	TGGCATGAAGGTTTCATCG
STM1586		6.6	-	13.5	1672845	1672862	TGGCAAGAATATTGCCAT
STM1594	<i>srfB</i>	4.8	+	13.4	1681565	1681582	TGGCACACGTTTTGCGCT
STM1665		4.2	-	11.7	1759185	1759168	TGGCATCATTTTTTCAAG
STM1904 <sup>^</sup>	<i>yecN</i>	3.6	+	5.9	1998988	1999005	TGGCAAACCTGTGGTATA
STM1928	<i>otsA</i>	5.3	+	8.1	2023398	2023381	TGGCAGGAGCGTTTTATT
STM1990	<i>yedA</i>	5.6	+	14.4	2072998	2073015	TGGCGCGCTTTTTGCCTT
STM2033	<i>cbiC</i>	4.4	-	2.2	2111221	2111238	CGGTATAAATAATGCACG
STM2115	<i>wcaA</i>	4.3	-	9.5	2198775	2198792	TGGCATATAAATTGAGAT
STM2181	<i>yohJ</i>	15	+	4.9	2277993	2278010	AGGCATTTTTCTTGCATC
STM2430	<i>cysK</i>	4.8	-	11.7	2544207	2544190	TGGCATCACTGTTGCAGT
STM2475		9.0	-	1.0	2585621	2585638	TGGCACATCAGGCAAAG
STM2476	<i>ypfG</i>	3.1	+	12.7	2586874	2586857	TGGCAGGTCACCTGCAAT
STM2517	<i>sinH</i>	4.6	-	11.1	2650462	2650479	TGGTACGGATCTTGCCAT
STM2563	<i>yfhG</i>	4.7	-	6.8	2705786	2705803	CGGCGTAATTTTTGCATC
STM2939	<i>ygcH</i>	10	+	10.9	3080061	3080044	CGGCACAGCTCTTGCATC
STM2957	<i>rumA</i>	5.5	+	14.5	3105809	3105792	TGGAACGCTTTTCGATT
STM3072		4.1	-	6.9	3234181	3234164	TGGCCATTGAATGCATC
STM3302	<i>yhbE</i>	5.5	+	12.6	3472042	3472025	TGGCATGATGGTCCAG
STM3535	<i>glgA</i>	8.0	+	11.6	3702315	3702298	AGGCATGTTTTATGAAA
STM3721	<i>rfaP</i>	13.5	+	8.3	3916283	3916300	TGGTACGTAATGCACG
STM3863 <sup>^</sup>		7.5	+	11.1	4072959	4072942	TGGCGGATTATTGCCAG
STM3919	<i>wzzE</i>	4.2	+	11.0	4128295	4128312	TGGCCTGCTATTGCCCT
STM3924	<i>wecD</i>	22	+	11.8	4133232	4133249	TGGCGCGAAATTGCACA
STM4013.S		3.6	-	13.6	4222708	4222725	TGGCATAAACCTGAAAA
STM4226	<i>yjbA</i>	6.4	-	4.2	4446318	4446301	AGGCGGAATAATGCATC
STM4290	<i>proP</i>	13	+	10.1	4532022	4532039	TGGCCTGATTTTTGCAGG
STM4572	<i>stjB</i>	8.3	-	8.2	4826908	4826925	TGGCGTGCGGATTCAAT

<sup>a</sup>Loci listed in **bold** are known or predicted  $\sigma^{54}$ -dependent promoters.

<sup>b</sup>Ratio of signals between WT and  $\Delta$ *rpoN* cells.

<sup>c</sup>Orientation of the predicted binding site with respect to the listed ORF. (+) binding site is in *same* direction as ORF; (-) binding site is in *opposite* direction as ORF.

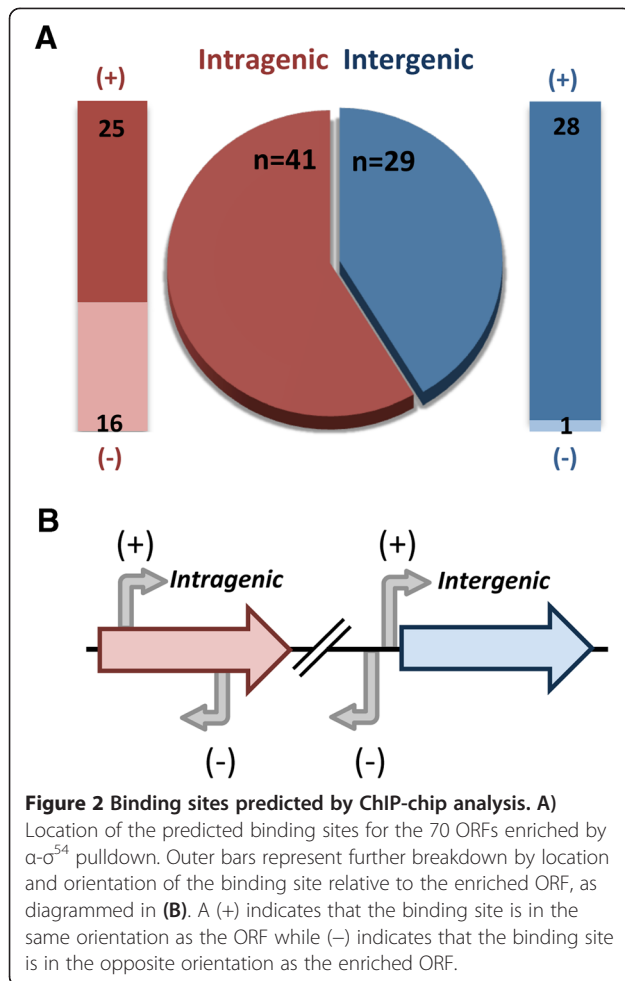
<sup>d</sup>PSSM score for the best predicted binding site within 1 kb of enriched ORF using the position-specific scoring matrix derived from the sequences in Additional file 2. As a reference to interpret the PSSM scores, the *S. Typhimurium* LT2 chromosome contains 3 sites with PSSM scores  $\geq 22.0$ , 21 sites with PSSM scores  $\geq 18.0$ , 61 sites with PSSM scores  $\geq 14.0$ , and 401 sites with PSSM scores  $\geq 10.0$ .

<sup>e</sup>The predicted binding site is a potential promoter for a neighboring gene based on its orientation and location within 250 bp of the 5' end of a neighboring gene or a long intergenic region (>100 bp) that may encode a sRNA.

intergenic sequence was identified by Motif Locator and is reported in Table 3. For the 70 enriched ORFs, 29 of the associated binding sites mapped to intergenic regions and 41 of the potential  $\sigma^{54}$  binding sites were located within the enriched ORF (Figure 2). In determining the most likely binding site for an enriched ORF, sequence within an adjacent non-enriched ORF was not considered for potential  $\sigma^{54}$  binding sites, since the ORF containing the binding site should be enriched; therefore, even if a site with a higher PSSM score was located in an immediately adjacent non-enriched ORF, the next highest scoring site found in either the enriched ORF or

adjacent intergenic sequence was reported as the potential binding site in Table 3. This reflects a limitation of the *in silico* prediction of  $\sigma^{54}$  binding sites based on a PSSM that was created with known and predicted intergenic promoter sequences; the sequences for intragenic promoters or for  $\sigma^{54}$  binding sites that are regulatory sites, but not promoters, may differ enough to appreciably affect PSSM scores.

Consensus sequences were generated using WebLogo [47] for the intergenic and intragenic potential  $\sigma^{54}$  binding sequences and for the promoter sequences used to generate the PSSM (Figure 3). Noteworthy differences in



the consensus sequence for the intragenic  $\sigma^{54}$  binding sites, as compared to the consensus sequences for the intergenic  $\sigma^{54}$  binding sites and PSSM promoters, are at the -23 and -11 positions, which each contribute in different ways to  $\sigma^{54}$ -promoter DNA interactions. The -23 A-T base pair is important in promoter recognition by  $\sigma^{54}$ ; the winged helix-turn-helix DNA binding motif of  $\sigma^{54}$  makes base-specific contacts with the top strand GG at positions -26 and -25 and with the bottom strand T at position -23 [48]. The base pairs immediately adjacent to the conserved GC element in the -12 region of the promoter are involved in  $E\sigma^{54}$  binding to form the stable closed complex; the bases on the bottom strand of the promoter at the -12 and -11 positions interact with  $\sigma^{54}$  in a short region of 'early melting' that stabilizes closed complex until the bEBP binds  $\sigma^{54}$  and activates the holoenzyme to transition to open complex [49]. The reduced conservation of nucleotide sequence at the -23 and -11 positions for the potential intragenic  $\sigma^{54}$  binding sites may reflect varied functionality of these intragenic sites, or a level of inaccuracy inherent to *in silico* prediction

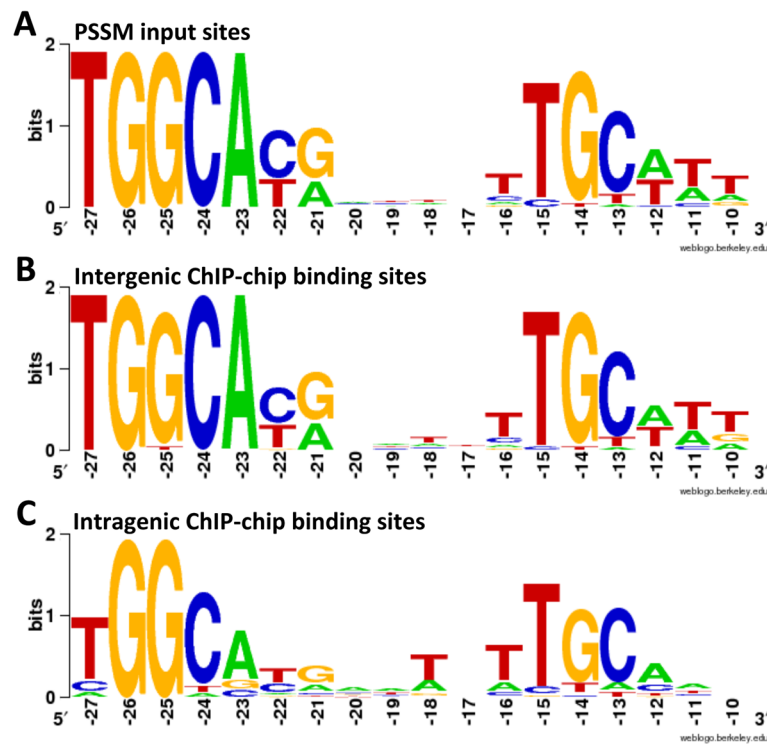
of the binding sites associated with enriched ORFs in the ChIP-chip assays.

The position and orientation for each potential  $\sigma^{54}$  binding site are indicated in Table 3 and summarized for all the binding sites in Figure 2. This information is useful in considering possible functions for the binding sites. For example, the 16 intragenic  $\sigma^{54}$  binding sites oriented in the opposing direction of the gene might regulate by transcription interference and/or anti-sense RNA [19]. Four intragenic  $\sigma^{54}$  binding sites are within 250 bp of the 5' end of a downstream gene, or a large intergenic region (>100 bp), and oriented in the direction such that they might act as promoters for the downstream gene or a sRNA [20]. Binding sites located near a functional  $\sigma^{54}$  promoter may serve to accelerate the search for the promoter by  $E\sigma^{54}$  sliding from the secondary sites [50]; while binding sites adjacent to, or overlapping, a  $\sigma^{70}$ - or  $\sigma^{54}$ -promoter may bind  $E\sigma^{54}$  or  $\sigma^{54}$  and repress or activate transcription from the other promoter [51]. The possible functions of the  $\sigma^{54}$  binding sites are quite varied and many are dependent on whether the binding site can function as a promoter.

It is likely that our initial approach to defining the global binding sites of  $\sigma^{54}$  in *S. Typhimurium* LT2 resulted in an underestimation of the number of binding sites. Multiple sites within ~2,000 bp of an ORF would enrich one or two adjacent ORFs and, in our analysis, would have been counted as one site. In addition, since  $E\sigma^{54}$ -promoter closed complexes are reversible [52], some complexes might have not been detected due to high disassociation rates; detection of these sites may be improved in the presence of DctD250, which stimulates conversion of closed complex to the more stable open complex, and rifampicin, which prevents extension of RNA past the second or third nucleotide [53], thus improving the chances of cross-linking  $E\sigma^{54}$  at the promoter sequence [54].

#### Promoter-reporter analysis to determine activity for predicted promoter sequences

To assess the functionality of  $\sigma^{54}$  binding sites defined by the ChIP-chip and PSSM analyses, promoter-*lacZ* fusion assays were performed using several of these sequences. We had two goals in performing these assays. First, we wanted to further confirm the  $\sigma^{54}$ -dependent promoter activity for some of the predicted  $\sigma^{54}$ -dependent promoters that were up-regulated in the DNA microarrays and enriched in the ChIP-chip assays. Secondly, we wanted to test the  $\sigma^{54}$  binding site predictions from the ChIP-chip combined with PSSM analyses; i.e. does a potential  $\sigma^{54}$  binding site equate to a  $\sigma^{54}$ -dependent promoter? This promoter function assay is the initial exploration of the roles for  $\sigma^{54}$  binding sites located within intragenic regions.



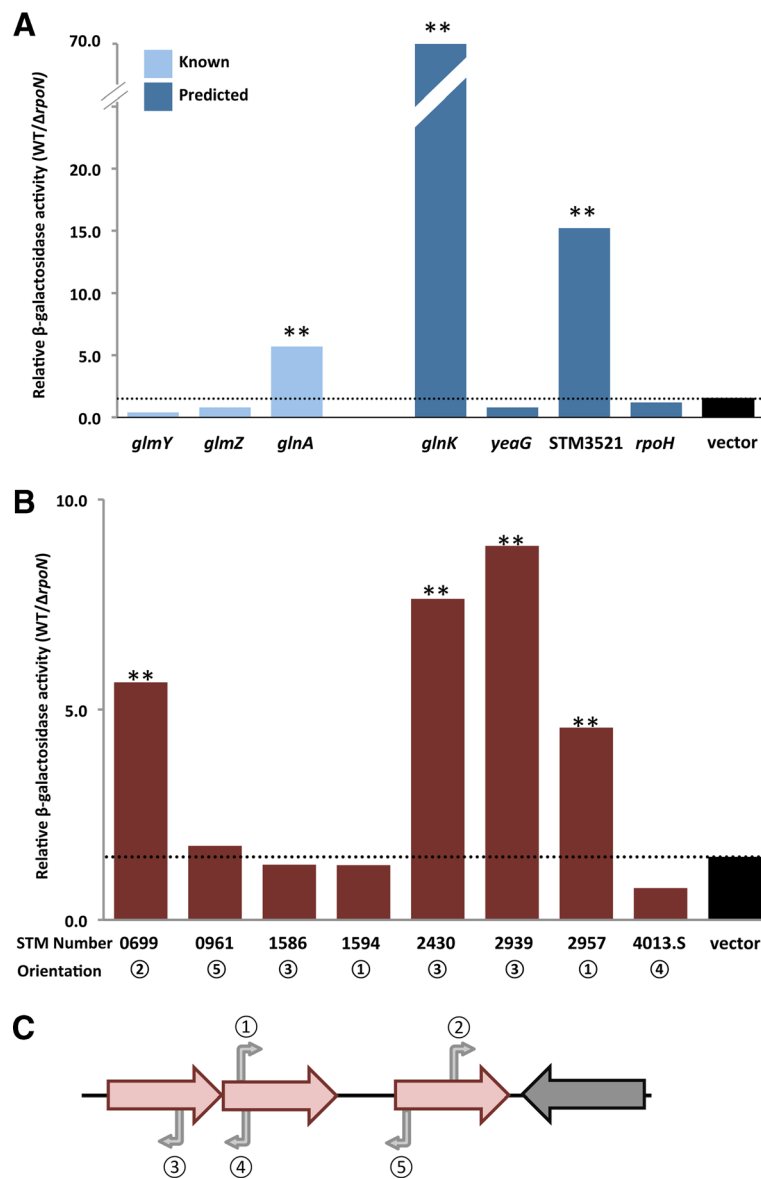
**Figure 3 Alignment of  $\sigma^{54}$  binding sites.** Weblogos show the consensus sequence for **A)** 27 known/predicted promoter sequences used to generate the position-specific scoring matrix **B)** 29 predicted intergenic binding sites for ORFs enriched in ChIP-chip analysis or **C)** 41 predicted intragenic binding sites from within ORFs enriched in ChIP-chip analysis. Weblogos were generated using the online program available at <http://weblogo.berkeley.edu/>.

Potential promoters were introduced upstream of a promoter-less *lacZ* gene in a reporter vector, either pDS11 or pDS12 (which differ only in their MCS sequence). These promoter-reporter plasmids were co-transformed along with the DctD250 expression plasmid into either WT or  $\Delta rpoN$  cells. After induction of DctD250 expression, standard  $\beta$ -galactosidase assays were performed. The results from WT were compared to those from the  $\Delta rpoN$  mutant to determine whether activity seen was  $\sigma^{54}$ -dependent (Figure 4). For intergenic sequences that were known or predicted to be  $\sigma^{54}$ -dependent promoters, the results matched those observed in the DNA microarray assays (Figure 4A, Table 2). The *glnA*, *glnK*, and STM3521 (*rtcBA* operon) promoters showed strong  $\sigma^{54}$ -dependent activity. For the *glmY*, *glmZ*, *rpoH*, and *yeaG* promoters, transcription in the  $\Delta rpoN$  mutant was either as high as or higher than in wild type cells. This is likely due to the presence of  $\sigma^{70}$ -type promoters in the cloned sequence. In addition to the  $\sigma^{54}$  dependent promoters, other promoters have been reported upstream of *glmY*, *glmZ*, *rpoH*, and *yeaG* [28,36,41].

A total of eight intragenic sites identified in the ChIP-chip assay were selected for functional analysis (Figure 4B). All of these predicted sites had PSSM scores  $>10$ . As shown in Figure 4C, the sites chosen represent a variety of

configurations with regard to their position and orientation within the ORF as well as the position and orientation of downstream ORFs. Given the possible functions for an intragenic promoter sequence (e.g. promoter for a downstream gene or sRNA, generation of antisense RNA, etc.), results of our analysis allow us to determine which, if any, of these roles may be attributable to any of these promoters.

Comparing the levels of *lacZ* expression in wild type cells to those in  $\Delta rpoN$  mutants, we found that four of the eight intragenic sites were able to function as  $\sigma^{54}$ -dependent promoters. For these sites, the difference observed between WT and  $\Delta rpoN$  cells varied from 4.6-fold for the sequence located in STM2957 to 8.9-fold for the sequence within STM2939. Overall, the level of transcription from these promoters was relatively low, with Miller units ranging from  $\sim 30$ -100. The low activity levels may indicate that  $E\sigma^{54}$  has a low affinity for these sequences or that the DctD250 is inefficient in productively engaging closed complexes formed at these sites. A subset of the promoter-reporter plasmids with intragenic sites that exhibited  $\sigma^{54}$ -dependent transcription were also assayed in WT cells versus WT-DctD250 to determine the dependence of transcription on the promiscuous, constitutive bEBP (Table 1). All three intragenic promoters assayed, STM0699, STM2430, and STM2939, gave low



**Figure 4 Promoter location, orientation, and activity for selected  $\sigma^{54}$  binding sites.** The ratio of  $\beta$ -galactosidase activity (Miller Units) in WT+DctD250 vs.  $\Delta$ rpoN+DctD250 cells is shown for (A) known (light blue bars) and predicted (dark blue bars)  $\sigma^{54}$ -dependent promoters, and (B) potential intragenic promoter sequences (red bars) in the promoter reporter vectors, pDS11 or pDS12 (black bars). Double asterisks denote significant increase in  $\beta$ -galactosidase activity in WT+DctD250 versus  $\Delta$ rpoN+DctD250 (p-value <0.02). Circled numbers below locus tags indicate orientation of the potential promoter sequence, as illustrated in (C). Orientation of potential intragenic promoter sequence is: **1**) same as ORF and >300 bp from 3' end; **2**) same as ORF and <300 bp from 3' end of a convergent downstream gene; **3**) opposite of ORF and >300 bp from 5' end; **4**) opposite of ORF and <300 bp from the 3' end of an upstream gene; and **5**) opposite of ORF and <300 bp from 5' end of gene, but >300 bp from the 3' end of an upstream gene.

levels of  $\beta$ -galactosidase activity in the absence of DctD250 and from 4.7- to 34-fold higher levels of  $\beta$ -galactosidase activity in the presence of DctD250. The  $\sigma^{54}$ -dependent transcription from intragenic binding sites suggests previously unrecognized regulatory functions for  $\sigma^{54}$  in *Salmonella*; however, it will be critical to characterize transcription from their chromosomal loci before biological functions can be ascribed.

Some of the potential promoter sequences that were assayed failed to show any transcriptional activity. There are a number of possible reasons for the lack of promoter activity for these sites. Two likely explanations are: 1) the wrong sequence was chosen as the binding site based on the PSSM score and proximity to the enriched ORF in the CHIP-chip assays, i.e. a lower scoring sequence near the enriched ORF was the actual  $\sigma^{54}$

binding site; or 2) the  $\sigma^{54}$  binding site does not function as a promoter but serves another regulatory role, such as an operator site for regulating promoter activity, a site for transient binding in facilitated diffusion, or a site for sequestering  $E\sigma^{54}$  in order to increase local concentration (since  $\sigma^{70}$  has a higher affinity for RNAP [55]).

#### Summary of *S. Typhimurium* LT2 $\sigma^{54}$ regulon and comparison to $\sigma^{54}$ regulons of other bacteria

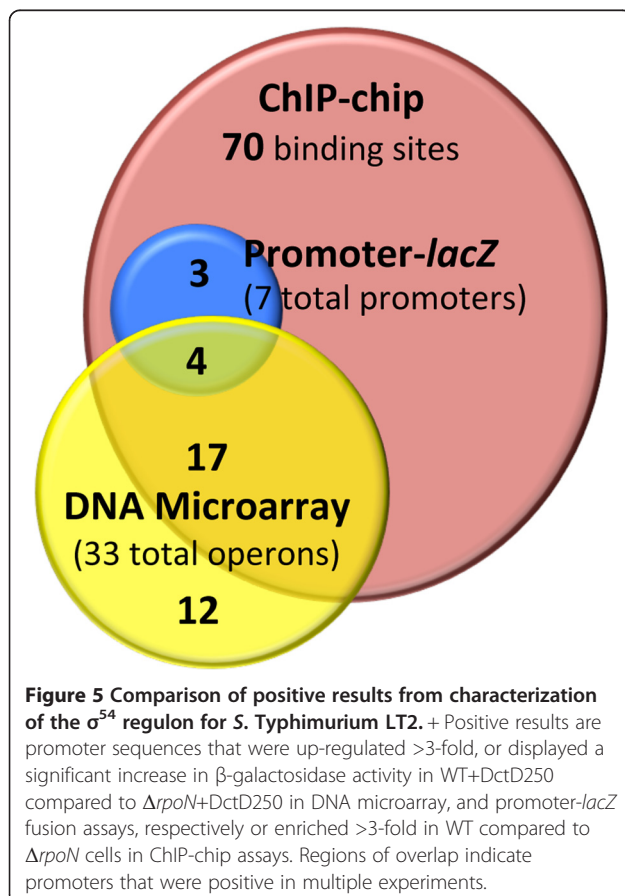
Figure 5 summarizes the results from the DNA microarray and promoter-fusion assays performed in the presence of DctD250 and ChIP-chip in the absence of DctD250 to characterize the  $\sigma^{54}$  regulon of *S. Typhimurium* LT2. Based on DNA microarray, there are 33 up-regulated operons (76 genes; Additional file 1); global ChIP-chip combined with *in silico* analysis revealed at least 70  $\sigma^{54}$  binding sites (Table 3), of which 21 were associated with up-regulated operons from the DNA microarrays. The promoter-*lacZ* fusions with seven of the 29 intergenic  $\sigma^{54}$  binding sites and eight of the 41 intragenic  $\sigma^{54}$  binding sites showed DctD250- and  $\sigma^{54}$ -dependent expression for three intergenic sites (associated with up-regulated operons) and four intragenic  $\sigma^{54}$  binding sites (Table 1, Figure 4). The cellular functions impacted by genes in the  $\sigma^{54}$  regulon of *S. Typhimurium* LT2 are quite

diverse, ranging from carbon-source and amino acid metabolism to response to stressors, such as nitric oxide and toxic levels of zinc (Table 2). Our results suggest that a new cellular process may be added to this extensive list—cell immunity through the CRISPR system; the role of  $\sigma^{54}$  in regulating a *casI*-related gene within an operon of CRISPR-associated genes is presently being investigated.

The  $\sigma^{54}$  global regulon of *S. Typhimurium* LT2 may differ from that of virulent *S. Typhimurium* isolates due to accumulated mutations in this extensively-used, laboratory strain, particularly the *rpoS* mutation that contributes to attenuation of the LT2 strain [42]. Changes in the level of expression of one sigma factor can alter the expression of genes that are expressed by different sigma factors [56]; for example, it has been shown that deletion of *rpoN* alters expression of  $\sigma^S$ -dependent promoters in *E. coli* [57]. We are currently characterizing the  $\sigma^{54}$  global regulon of the virulent strain *S. Typhimurium* 14028s.

The  $\sigma^{54}$  regulons in other  $\delta/\gamma$ -proteobacteria have been characterized experimentally to varying extents [13-15,58-60]. Only in *Vibrio cholera* 037 strain V52 have both global transcripts and binding sites been characterized experimentally [14]. In *E. coli* MG1655 and *Geobacter sulfurreducens*, the global  $\sigma^{54}$  transcriptomes were determined and local  $\sigma^{54}$  binding sites associated with up-regulated genes were assessed by computational analysis and selected promoters were assessed experimentally [13,15]. The number and diversity of the operons that are directly controlled by  $\sigma^{54}$ -promoters in these  $\delta/\gamma$ -proteobacteria are comparable to that of *S. Typhimurium* LT2. The greatest variability in the  $\sigma^{54}$  regulons of the  $\gamma$ -proteobacteria appears to be the location of  $\sigma^{54}$  binding sites. Zhao et al. [13] estimated 70  $\sigma^{54}$  promoters in *E. coli* MG1655, of which 13 (18%) were intragenic or located between convergently transcribed genes. In *V. cholera*, Dong and Mekalanos [14] identified a total 68  $\sigma^{54}$  binding sites, of which 35 (51%) were intragenic and, similarly, we found 70 potential  $\sigma^{54}$  binding sites of which 41 (58%) appear to be located in intragenic regions.

Does the success with DctD250 in characterizing the *S. Typhimurium*  $\sigma^{54}$  regulon predict utility of this constitutive, promiscuous activator in defining  $\sigma^{54}$  global regulons in bacteria from other classes in the Proteobacteria phylum, or from other phyla? The key to activation of  $E\sigma^{54}$  by DctD250 in diverse bacteria is the ability of the activator to make the appropriate interactions with  $\sigma^{54}$  in the context of the  $E\sigma^{54}$ -promoter closed complex; thus, comparison of interacting regions of  $\sigma^{54}$  and bEBPs between *S. Typhimurium* and phylogenetically diverse bacteria is a good predictor of success. Extensive characterization of bEBP activation of  $E\sigma^{54}$  in



closed complex has shown that the GAFTGA motif of the AAA+ ATPase domain plays a primary and essential role for productive interactions with  $\sigma^{54}$ , which lead to transcriptional activation (reviewed in [8]); the GAFTGA motif is very highly conserved among bEBPs in all bacteria that encode  $\sigma^{54}$ , which includes bacteria from a majority of the eubacterial phyla [61]. It has not yet been determined which specific residues of  $\sigma^{54}$  are contacted by Loop 1 of the bEBP AAA+ ATPase domain, but it has been clearly demonstrated that multiple residues within the amino-terminal 50 amino acids of  $\sigma^{54}$  (Region I) are key determinants for activator interaction [62] and there is extensive conservation of amino acid sequence in Region I for  $\sigma^{54}$  from phylogenetically diverse bacteria [63]. Thus, the comparison of interacting regions of  $\sigma^{54}$  and the AAA+ ATPase domain among diverse bacteria predicts that DctD250 will be a valuable tool in characterizing the  $\sigma^{54}$  regulons in many bacteria.

## Conclusions

The results of DNA microarray and promoter-*lacZ* fusion analyses of the  $\sigma^{54}$  regulon of *S. Typhimurium* LT2 in the presence of DctD250 support our initial hypothesis: the AAA+ ATPase activation domain of DctD can stimulate transcription from  $\sigma^{54}$ -dependent promoters in a constitutive and promiscuous manner, thereby facilitating the global characterization of  $\sigma^{54}$  regulons. Sixteen previously predicted  $\sigma^{54}$ -dependent operons were confirmed, and a new  $\sigma^{54}$ -dependent gene, *casI*, was identified by the DNA microarray and ChIP-chip analyses. In addition, the ChIP-chip analyses indicate an excess of  $\sigma^{54}$  binding sites compared to the number of  $\sigma^{54}$ -dependent transcripts and a high percentage of intragenic binding sites, suggesting that  $\sigma^{54}$  and  $\sigma^{54}$  may have more regulatory functions than transcription initiation at the start of an operon or sRNA. The number of functional promoters located inside genes suggests a need to consider such promoters in bioinformatic analyses of transcription factor binding sites.

## Methods

### Bacterial strains, media, and enzymes

The parental strain, designated wild-type, in these experiments was *Salmonella enterica subspecies enterica serovar Typhimurium* LT2 derivative MS1868 [*leuA414* (Am) *hdsSB*(r<sup>m+</sup>)*Fels*] [64]. An isogenic derivative, TRH134, has a deletion in *rpoN* (*ntrA*) from codons 8 through 455, rendering it auxotrophic for glutamine [65]. *S. Typhimurium* strains were cultured in either nutrient broth (NB; Difco Laboratories), MOPS minimal media [66], or nitrogen-limiting MOPS [67]. Media supplement concentrations were 5 mM L-glutamine (Gln), 40  $\mu$ g/ml L-Leucine (Leu), and 10 mM L-glutamate (Glu). Cloning procedures were performed in *E. coli* DH5 $\alpha$  cultured in Luria-Bertani medium (LB; Fisher

Scientific). All strains were grown at 37°C. Antibiotics (Sigma-Aldrich) were used at the following concentrations ( $\mu$ g/ml) for *E. coli*/*S. Typhimurium* (NB)/*S. Typhimurium* (MOPS), respectively: ampicillin (Amp) 80/120/50; spectinomycin (Spc) 50/125/50; streptomycin (Str) 25/75/0. All enzymes were purchased from New England Biolabs, unless otherwise indicated, and were used according to manufacturer's recommendations.

### Plasmids

Plasmid pPBHP92 is a derivative of the expression vector pTrcHisC (Invitrogen) that expresses the *Sinorhizbium meliloti* DctD AAA+ ATPase domain (E141-S390, designated DctD250) with an N-terminal 6x-His tag. This plasmid was constructed by digestion of pHX182 [17] with NdeI, filling in the 5'-overhang with T4 DNA polymerase and subsequent digestion with XhoI. The blunt-XhoI fragment containing the truncated *dctD* was cloned into pTrcHisC, which had been cut with NheI, blunt-ended, and cut with XhoI. The truncated *dctD* is under control of  $P_{trc}$  and subject to repression by the vector-encoded LacI. The reporter plasmids used in these studies, pDS11 and pDS12, are both derivatives of pDV6 [25] that contain a promoter-less copy of *lacZ* downstream of a MCS region. The MCS region was generated by annealing two oligonucleotide primers (Additional file 3) which were then ligated into a pDV6 backbone that had been digested with BamHI and HindIII. pDS11 and pDS12 differ only in MCS sequence. Potential promoter sequences were amplified from *S. Typhimurium* LT2 genomic DNA using Taq polymerase and the primers in (Additional file 3) and cloned into pCR2.1 (Invitrogen). Sequencing analysis to determine accuracy and orientation was performed for all plasmids by Genewiz, Inc. (South Plainfield, NJ). Depending on their orientation in pCR2.1 potential promoter sequences were sub-cloned into pDS11/12 using XbaI and either KpnI or HindIII. Plasmid pTG4, which encodes the DctD AAA+ ATPase domain under control of  $P_{tac}/lacI^q$ , was created by amplifying the corresponding region of pPBHP92 using primers DctD-F/R (Additional file 3), digesting the product with BamHI and HindIII, followed by ligation into the similarly digested pKH66 [68].

### Transcriptional profiling by microarrays

*S. Typhimurium* strains MS1868 and TRH134, each bearing plasmid pPBHP92 (WT+DctD250 and  $\Delta$ *rpoN*+DctD250, respectively), were grown overnight at 37°C in NB-Amp. Cultures were sub-cultured in fresh medium and grown to mid-log phase ( $OD_{600} \approx 0.8$ ). Since the basal level of DctD250 expression from pPBHP92 was shown to optimally activate transcription from a  $\sigma^{54}$ -dependent *dctA'*-*lacZ* reporter [17], IPTG induction was not used for these cultures. RNA isolated using the RNAeasy kit (Qiagen) was used to generate differentially

labeled cDNA using reverse transcriptase as previously described [69]. Labeled cDNA was hybridized to DNA microarrays containing complete open reading frames (ORFs) from *S. Typhimurium* LT2 printed in triplicate [70]. Microarrays were scanned with a ScanArray Lite laser scanner (Packard BioChip Technologies, Billerica, MA) using ScanArray Express 1.1 software. Signal intensities were quantified using QuantArray 3.0 (Packard). The ratio of WT+DctD250 signal to  $\Delta rpoN$ +DctD250 signal was determined for each of the triplicate spots and the median value for each ORF was used in the statistical analysis [70]. Data shown is the result of three biological replicates with statistical analysis performed using the WebArrayDB program [71,72]. The intensity values for the three biological replicates of WT+DctD250 and of  $\Delta rpoN$ +DctD250 were compared for the calculation of the p-values, where the null hypothesis was that the intensities for WT+DctD250 and  $\Delta rpoN$ +DctD250 would be equivalent. Genes that displayed a WT+DctD250/ $\Delta rpoN$ +DctD250 signal ratio of >3-fold with a p-value of <0.02 were considered to be up-regulated.

#### Chromatin immunoprecipitation (ChIP)

ChIP was carried out using the ChIP Assay kit (USB Corporation) essentially as described by the manufacturer's instructions. Briefly, 100 ml cultures of *S. Typhimurium* strains MS1868 (WT) and TRH134 ( $\Delta rpoN$ ) were grown overnight in NB at 37°C and sub-cultured in fresh medium the next day. Once cultures reached mid-log phase ( $OD_{600} \approx 0.7$ ), cells were treated with formaldehyde (3 ml of a 37% solution per 100 ml of culture) for 10 min. at room temperature to cross-link proteins to DNA. Cross-linking was quenched by the addition of glycine (10 ml of 1.33 M solution per 100 ml of culture) and incubation at 4°C for 30 min. Cells were harvested, washed and lysed in accordance with kit instructions. Cells were lysed in two passages through a French pressure cell at 10,000 psi. Cell extracts were clarified and pre-cleared with the provided protein A-Sepharose bead slurry per the kit instructions. 0.6 ml of the resulting extracts were mixed with 2  $\mu$ l of rabbit anti-serum against *S. Typhimurium*  $\sigma^{54}$  [73] and incubated with gentle shaking overnight at 4°C. The next day, 50  $\mu$ l of protein A-Sepharose bead slurry was added to each sample, incubated 1 hr at room temperature and collected by centrifugation. The beads were washed, and protein-DNA complexes were eluted from the beads and disrupted per the supplier's instructions. DNA was purified from each sample using the Qiagen PCR purification kit.

#### ChIP-chip assays

Purified ChIP DNA was amplified by ligation-mediated PCR, adapting the procedure found at [[http://www.flychip.org.uk/protocols/archive\\_protocols/lm\\_pcr.php](http://www.flychip.org.uk/protocols/archive_protocols/lm_pcr.php)].

Linkers consisting of complementary oligonucleotides (LM-PCR; Additional file 3) were ligated to the ends of purified DNA repaired with T4 DNA polymerase. Ligated was purified using the Qiagen PCR purification kit and the DNA was amplified with Taq polymerase (Fermentas; Burlington, ON) using LM-PCR-R as the PCR primer and the following cycling conditions: 55°C—2 min (1 $\times$ ); 72°C—5 min (1 $\times$ ); 94°C—5 min (1 $\times$ ); 94°C—1 min, 55°C—1 min, 72°C—1 min (24 $\times$ ); 72°C—5 min (1 $\times$ ); 4°C—hold. The resulting amplicons, most of which were 300–800 bp, were purified using the Qiagen PCR purification kit and to prepare dye-labeled DNA (Cy3 or Cy5) for hybridization to the *S. Typhimurium* complete ORF microarray. Microarrays were scanned and analyzed as above. ChIP-chip was performed on three biological replicates for the WT and  $\Delta rpoN$  strains; the statistical analysis of the data was performed as described for the microarray data.

#### Identifying candidate $\sigma^{54}$ binding sites in the *S. Typhimurium* genome

The Motif Locator program [<http://www.cmbi.uga.edu/software.html>] was used to identify candidate  $\sigma^{54}$  binding sites. The program applies the standard position-specific score matrix (PSSM) described in [46]. We used a PSSM derived from the alignment of 27 high-confidence sites supported by experimental evidence in either *Salmonella* or *E. coli* (Additional file 2). Background nucleotide frequencies were assigned in accordance with the genomic G+C content. Pseudo-counts equal to the background frequencies were used in PSSM construction. For ORFs discovered in the ChIP-chip assay, this matrix was used to determine the most likely binding site either within the ORF itself or in the region  $\pm 200$ , 500, or 1000 bp surrounding the gene.

#### $\beta$ -Galactosidase assays

The DctD250 expression plasmid pTG4 was introduced into *S. Typhimurium* MS1868 and TRH134 by electroporation using a GenePulser 2 system (BioRad; Hercules, CA) and the resulting transformants were electroporated with pDS11 or pDS12 reporter constructs containing potential  $\sigma^{54}$ -dependent promoter sequences. Overnight cultures grown in MOPS-LeuGln, or nitrogen-limiting MOPS-Glu, with the appropriate antibiotics were sub-cultured into fresh medium, grown to  $OD_{600} \approx 0.2$ , and induced with 50  $\mu$ M IPTG (empirically determined IPTG concentration for optimal expression of DctD250 from pTG4 to activate known  $\sigma^{54}$ -dependent promoters on the reporter plasmids). Cultures were induced for 6 hours and  $\beta$ -galactosidase activity was measured as described previously [74] with the following changes: 1) assays were performed at 37°C, and 2) after stopping



reaction, samples were centrifuged and OD<sub>420</sub> of the supernatant was measured, eliminating the OD<sub>550</sub> correction for cell debris. Activity was calculated as Miller units: [(OD<sub>420</sub> × 1000)/(OD<sub>600</sub> × Time(min) × volume (ml))] [74]. Ratios of activity in wild type/ $\Delta$ *rpoN* cells were compared and analyzed using a 2-tailed Student's T-test. Data shown for each promoter construct represents  $\geq 3$  biological replicates.

#### Accession number for microarray and ChIP-chip data

The DNA microarray and ChIP-chip data were deposited in NCBI GEO under accession number GSE25849.

#### Additional files

**Additional file 1: Complete list of microarray results.** The file contains data for all genes within operons that showed up-regulation by  $\sigma^{54}$ .

**Additional file 2: Sequences used to generate Position-Specific Scoring Matrix.**

**Additional file 3: Oligonucleotides used.**

#### Abbreviations

Amp: Ampicillin; bEBP: Bacterial enhancer-binding protein; ChIP: Chromatin immunoprecipitation; E $\sigma$ : RNA polymerase holoenzyme; LB: Luria-bertani media; LM-PCR: Ligation-mediated PCR; NB: Nutrient broth; RNAP: RNA polymerase; Spc: Spectinomycin; Str: Streptomycin; UAS: Upstream activation sequence.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

DS harvested RNA, carried out promoter-reporter assays, and drafted and revised and prepared the manuscript. JF prepared cDNA and preformed DNA microarrays assays and applied DNA from the ChIP to the microarrays. SP and MM produced the microarrays and protocols and performed the statistical analysis of the microarray therein. JM performed the bioinformatic analyses. TH conceived, designed, and coordinated the study, harvested RNA, and performed ChIP pulldowns. AK conceived, designed, and coordinated the study, harvested RNA, and drafted and revised the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

We would like to thank the following people for their assistance on this project: Ashley Bono, Uchenna Ewulonu, and Trevor Wright generated reporter constructs. Tanya Grancharova sub-cloned DctD250. Sonya Chelliah performed ChIP pull-downs. Jennifer Turpin prepared and applied cDNA to the microarrays. This work was supported by National Science Foundation Grant MCB-1051175 (to T.R.H. and A.C.K.). MM and SP were supported in part by NIH Contract No. HHSN272200900040C from PATRIC and grants AI039557 AI052237, AI073971, AI075093, AI077645 AI083646, USDA grants 2009-03579 -30127 and 2011-67017-30127, the Binational Agricultural Research and Development Fund. JM was supported in part by the National Science Foundation grant DBI-0950266.

#### Author details

<sup>1</sup>Department of Microbiology, University of Georgia, 30602, Athens, GA, USA.

<sup>2</sup>US Department of Agriculture, Bacterial Epidemiology and Antimicrobial Resistance Research Unit, Richard B. Russell Research Center, Agricultural Research Service, 30605, Athens, GA, USA. <sup>3</sup>Department of Microbiology and Molecular Genetics, University of California, 92697, Irvine, CA, USA. <sup>4</sup>Institute of Bioinformatics, University of Georgia, 30602, Athens, GA, USA.

Received: 19 April 2013 Accepted: 28 August 2013

Published: 5 September 2013

#### References

1. Gruber TM, Gross CA: Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* 2003, **57**:441–466.
2. Ferro-Luzzi Ames G, Nikaïdo K: Nitrogen regulation in *Salmonella typhimurium*. Identification of an NtrC protein-binding site and definition of a consensus binding sequence. *EMBO J* 1985, **4**(2):539–547.
3. Leonhartsberger S, Huber A, Lottspeich F, Bock A: The *hydH/G* Genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system. *J Mol Biol* 2001, **307**(1):93–105.
4. Palacios S, Escalante-Semerena JC: *prpR*, *ntrA*, and *ihf* functions are required for expression of the *prpBCDE* operon, encoding enzymes that catabolize propionate in *Salmonella enterica* serovar Typhimurium LT2. *J Bacteriol* 2000, **182**(4):905–910.
5. Weiner L, Brisette JL, Model P: Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on sigma<sup>54</sup> and modulated by positive and negative feedback mechanisms. *Genes Dev* 1991, **5**(10):1912–1923.
6. Niehus E, Gressmann H, Ye F, Schlapbach R, Dehio M, Dehio C, Stack A, Meyer TF, Suerbaum S, Josenhans C: Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*. *Mol Microbiol* 2004, **52**(4):947–961.
7. Wigneshwaraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, et al: Modus operandi of the bacterial RNA polymerase containing the  $\sigma^{54}$  promoter-specificity factor. *Mol Microbiol* 2008, **68**(3):538–546.
8. Bush M, Dixon R: The role of bacterial enhancer binding proteins as specialized activators of sigma<sup>54</sup>-dependent transcription. *Microbiol Mol Biol Rev* 2012, **76**(3):497–529.
9. Barrios H, Valderrama B, Morett E: Compilation and analysis of sigma (54)-dependent promoter sequences. *Nucleic Acids Res* 1999, **27**(22):4305–4313.
10. Ghosh T, Bose D, Zhang X: Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol Rev* 2010, **34**(5):611–627.
11. Belitsky BR, Sonenshein AL: An enhancer element located downstream of the major glutamate dehydrogenase gene of *Bacillus subtilis*. *Proc Natl Acad Sci USA* 1999, **96**(18):10290–10295.
12. Reitzer LJ, Magasanik B: Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. *Cell* 1986, **45**(6):785–792.
13. Zhao K, Liu M, Burgess RR: Promoter and regulon analysis of nitrogen assimilation factor, sigma<sup>54</sup>, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. *Nucleic Acids Res* 2010, **38**(4):1273–1283.
14. Dong TG, Mekalanos JJ: Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholerae* O37 strain V52. *Nucleic Acids Res* 2012, **40**(16):7766–7775.
15. Leang C, Krushkal J, Ueki T, Puljic M, Sun J, Juarez K, Nunez C, Reguera G, DiDonato R, Postier B, et al: Genome-wide analysis of the RpoN regulon in *Geobacter sulfurreducens*. *BMC Genomics* 2009, **10**:331.
16. Brahmachary P, Dashti MG, Olson JW, Hoover TR: *Helicobacter pylori* FlgR is an enhancer-independent activator of  $\sigma^{54}$ -RNA polymerase holoenzyme. *J Bacteriol* 2004, **186**(14):4535–4542.
17. Xu H, Gu B, Nixon BT, Hoover TR: Purification and characterization of the AAA+ domain of *Sinorhizobium meliloti* DctD, a sigma<sup>54</sup>-dependent transcriptional activator. *J Bacteriol* 2004, **186**(11):3499–3507.
18. Studholme DJ: Enhancer-dependent transcription in *Salmonella enterica* Typhimurium: new members of the sigma<sup>N</sup> regulon inferred from protein sequence homology and predicted promoter sites. *J Mol Microbiol Biotechnol* 2002, **4**(4):367–374.
19. Georg J, Hess WR: cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev* 2011, **75**(2):286–300.
20. Gottesman S, Storz G: Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harbor perspectives in biology* 2011, **3**:12.
21. Kroger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, et al: The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci USA* 2012, **109**(20):E1277–1286.
22. van Heeswijk WC, Hoving S, Molenaar D, Stegeman B, Kahn D, Westerhoff HV: An alternative PII protein in the regulation of glutamine synthetase in *Escherichia coli*. *Mol Microbiol* 1996, **21**(1):133–146.
23. Genschik P, Drabikowski K, Filipowicz W: Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma<sup>54</sup>-regulated operon. *J Biol Chem* 1998, **273**(39):25516–25526.

24. Huala E, Stigter J, Ausubel FM: The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. *J Bacteriol* 1992, **174**(4):1428–1431.
25. Perkins-Balding D, Duval-Valentin G, Glasgow AC: Excision of IS492 requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. *J Bacteriol* 1999, **181**(16):4937–4948.
26. Porwollik S, Frye J, Florea LD, Blackmer F, McClelland M: A non-redundant microarray of genes for two related bacteria. *Nucleic Acids Res* 2003, **31**(7):1869–1876.
27. Klose KE, Mekalanos JJ: Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates *Salmonella typhimurium* virulence. *Infect Immun* 1997, **65**(2):587–596.
28. Gopel Y, Luttmann D, Heroven AK, Reichenbach B, Dersch P, Gorke B: Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in Enterobacteriaceae. *Nucleic Acids Res* 2011, **39**(4):1294–1309.
29. Hirschman J, Wong PK, Sei K, Keener J, Kustu S: Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription in vitro: evidence that the *ntrA* product is a sigma factor. *Proc Natl Acad Sci USA* 1985, **82**(22):7525–7529.
30. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA, Kustu S: Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci USA* 2000, **97**(26):14674–14679.
31. Kiupakis AK, Reitzer L: ArgR-independent induction and ArgR-dependent superinduction of the *astCADBE* operon in *Escherichia coli*. *J Bacteriol* 2002, **184**(11):2940–2950.
32. Riordan JT, Tietjen JA, Walsh CW, Gustafson JE, Whittam TS: Inactivation of alternative sigma factor 54 (RpoN) leads to increased acid resistance, and alters locus of enterocyte effacement (LEE) expression in *Escherichia coli* O157: H7. *Microbiology* 2010, **156**(Pt 3):719–730.
33. Gardner AM, Gessner CR, Gardner PR: Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli*. Role of NorR and  $\sigma^{54}$  in the nitric oxide stress response. *J Biol Chem* 2003, **278**(12):10081–10086.
34. Maier T, Binder U, Bock A: Analysis of the *hydA* locus of *Escherichia coli*: two genes (*hydN* and *hypF*) involved in formate and hydrogen metabolism. *Arch Microbiol* 1996, **165**(5):333–341.
35. Lutz S, Bohm R, Beier A, Bock A: Characterization of divergent NtrA-dependent promoters in the anaerobically expressed gene cluster coding for hydrogenase 3 components of *Escherichia coli*. *Mol Microbiol* 1990, **4**(1):13–20.
36. Janaszak A, Majczak W, Nadratowska B, Szalewska-Palasz A, Konopa G, Taylor A: A sigma<sup>54</sup>-dependent promoter in the regulatory region of the *Escherichia coli* *rpoH* gene. *Microbiology* 2007, **153**(Pt 1):111–123.
37. Ramirez-Santos J, Collado-Vides J, Garcia-Varela M, Gomez-Eichelmann MC: Conserved regulatory elements of the promoter sequence of the gene *rpoH* of enteric bacteria. *Nucleic Acids Res* 2001, **29**(2):380–386.
38. Lloyd LJ, Jones SE, Jovanovic G, Gyaneshwar P, Rolfe MD, Thompson A, Hinton JC, Buck M: Identification of a new member of the phage shock protein response in *Escherichia coli*, the phage shock protein G (PspG). *J Biol Chem* 2004, **279**(53):55707–55714.
39. Birkmann A, Sawers RG, Bock A: Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*. *Mol Gen Genet* 1987, **210**(3):535–542.
40. Miller KA, Phillips RS, Mrazek J, Hoover TR: *Salmonella* utilizes D-glucosaminase via a mannose family phosphotransferase system permease and associated enzymes. *J Bacteriol* 2013, **195**(18):4057–4066.
41. Ibanez-Ruiz M, Robbe-Saule V, Hermant D, Labrude S, Norel F: Identification of RpoS (sigma<sup>S</sup>)-regulated genes in *Salmonella enterica* serovar typhimurium. *J Bacteriol* 2000, **182**(20):5749–5756.
42. Swords WE, Cannon WM, Benjamin WH Jr: Avirulence of LT2 strains of *Salmonella typhimurium* results from a defective *rpoS* gene. *Infect Immun* 1997, **65**(6):2451–2453.
43. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P: CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007, **315**(5819):1709–1712.
44. Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA: Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 2009, **17**(6):904–912.
45. Buck M, Cannon W: Specific binding of the transcription factor  $\sigma^{54}$  to promoter DNA. *Nature* 1992, **358**(6385):422–424.
46. Mrazek J: Finding sequence motifs in prokaryotic genomes—a brief practical guide for a microbiologist. *Brief Bioinform* 2009, **10**(5):525–536.
47. Crooks GE, Hon G, Chandonia JM, Brenner SE: WebLogo: a sequence logo generator. *Genome Res* 2004, **14**(6):1188–1190.
48. Doucleff M, Pelton JG, Lee PS, Nixon BT, Wemmer DE: Structural basis of DNA recognition by the alternative sigma-factor,  $\sigma^{54}$ . *J Mol Biol* 2007, **369**(4):1070–1078.
49. Cannon W, Wigneshweraraj SR, Buck M: Interactions of regulated and deregulated forms of the sigma<sup>54</sup> holoenzyme with heteroduplex promoter DNA. *Nucleic Acids Res* 2002, **30**(4):886–893.
50. Li GW, Elf J: Single molecule approaches to transcription factor kinetics in living cells. *FEBS Lett* 2009, **583**(24):3979–3983.
51. Johansson LU, Solera D, Bernardo LM, Moscoso JA, Shingler V: sigma<sup>54</sup>-RNA polymerase controls sigma<sup>70</sup>-dependent transcription from a non-overlapping divergent promoter. *Mol Microbiol* 2008, **70**(3):709–723.
52. Friedman LJ, Gelles J: Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell* 2012, **148**(4):679–689.
53. Feklistov A, Mekler V, Jiang Q, Westblade LF, Irschik H, Jansen R, Mustaev A, Darst SA, Ebricht RH: Rifamycins do not function by allosteric modulation of binding of Mg<sup>2+</sup> to the RNA polymerase active center. *Proc Natl Acad Sci USA* 2008, **105**(39):14820–14825.
54. Raffaele M, Kanin EI, Vogt J, Burgess RR, Ansari AZ: Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo. *Mol Cell* 2005, **20**(3):357–366.
55. Maeda H, Fujita N, Ishihama A: Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res* 2000, **28**(18):3497–3503.
56. Grigorova IL, Phleger NJ, Mutalik VK, Gross CA: Insights into transcriptional regulation and sigma competition from an equilibrium model of RNA polymerase binding to DNA. *Proc Natl Acad Sci USA* 2006, **103**(14):5332–5337.
57. Dong T, Yu R, Schellhorn H: Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*. *Mol Microbiol* 2011, **79**(2):375–386.
58. Chaudhuri RR, Yu L, Kanji A, Perkins TT, Gardner PP, Choudhary J, Maskell DJ, Grant AJ: Quantitative RNA-seq analysis of the *Campylobacter jejuni* transcriptome. *Microbiology* 2011, **157**(Pt 10):2922–2932.
59. Damron FH, Owings JP, Okkotsu Y, Varga JJ, Schurr JR, Goldberg JB, Schurr MJ, Yu HD: Analysis of the *Pseudomonas aeruginosa* regulon controlled by the sensor kinase KinB and sigma factor RpoN. *J Bacteriol* 2012, **194**(6):1317–1330.
60. Dasgupta N, Wolfgang MC, Goodman AL, Arora SK, Jyot J, Lory S, Ramphal R: A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*. *Mol Microbiol* 2003, **50**(3):809–824.
61. Francke C, Groot Kormelink T, Hagemeyer Y, Overmars L, Sluijter V, Moezelaar R, Siezen RJ: Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics* 2011, **12**:385.
62. Xiao Y, Wigneshweraraj SR, Weinzierl R, Wang YP, Buck M: Construction and functional analyses of a comprehensive sigma<sup>54</sup> site-directed mutant library using alanine-cysteine mutagenesis. *Nucleic Acids Res* 2009, **37**(13):4482–4497.
63. Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD: The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol* 2000, **182**(15):4129–4136.
64. Grana D, Youderian P, Susskind MM: Mutations that improve the ant promoter of *Salmonella* phage P22. *Genetics* 1985, **110**(1):1–16.
65. Kelly MT, Hoover TR: Mutant forms of *Salmonella typhimurium* sigma<sup>54</sup> defective in transcription initiation but not promoter binding activity. *J Bacteriol* 1999, **181**(11):3351–3357.
66. Maloy S, Stewart VJ, Taylor RK: *Genetic Analysis of Pathogenic Bacteria: A Laboratory Manual*. Plainview, NY: Cold Spring Harbor Laboratory Press; 1996.
67. Kustu SG, McFarland NC, Hui SP, Esmon B, Ames GF: Nitrogen control of *Salmonella typhimurium*: co-regulation of synthesis of glutamine synthetase and amino acid transport systems. *J Bacteriol* 1979, **138**(1):218–234.

68. Hughes KT, Youderian P, Simon MI: **Phase variation in *Salmonella*: analysis of Hin recombinase and hix recombination site interaction in vivo.** *Genes Dev* 1988, **2**(8):937–948.
69. Wang Q, Frye JG, McClelland M, Harshey RM: **Gene expression patterns during swarming in *Salmonella typhimurium*: genes specific to surface growth and putative new motility and pathogenicity genes.** *Mol Microbiol* 2004, **52**(1):169–187.
70. Porwollik S, McClelland M: **Lateral gene transfer in *Salmonella*.** *Microbes Infect* 2003, **5**(11):977–989.
71. Xia XQ, McClelland M, Porwollik S, Song W, Cong X, Wang Y: **WebArrayDB: cross-platform microarray data analysis and public data repository.** *Bioinformatics* 2009, **25**(18):2425–2429.
72. Xia X, McClelland M, Wang Y: **WebArray: an online platform for microarray data analysis.** *BMC Bioinforma* 2005, **6**:306.
73. Kelly MT, Hoover TR: **The amino terminus of *Salmonella enterica* serovar Typhimurium sigma<sup>54</sup> is required for interactions with an enhancer-binding protein and binding to fork junction DNA.** *J Bacteriol* 2000, **182**(2):513–517.
74. Miller JH: *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria.* Plainview, NY: Cold Spring Harbor Press; 1992.

doi:10.1186/1471-2164-14-602

**Cite this article as:** Samuels *et al.*: Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the  $\sigma^{54}$  (RpoN) regulon of *Salmonella* Typhimurium LT2. *BMC Genomics* 2013 **14**:602.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

