Published online 6 March 2013

Nucleic Acids Research, 2013, Vol. 41, No. 8 ***4549–4564***
doi:10.1093/nar/gkt148

A comparison of dense transposon insertion libraries in the *Salmonella* serovars Typhi and Typhimurium

Lars Barquist*, Gemma C. Langridge, Daniel J. Turner, Minh-Duy Phan, A. Keith Turner, Alex Bateman, Julian Parkhill, John Wain* and Paul P. Gardner

Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK

Received November 7, 2012; Revised and Accepted February 14, 2013

ABSTRACT

Salmonella **Typhi and Typhimurium diverged only ~50 000 years ago, yet have very different host ranges and pathogenicity. Despite the availability of multiple whole-genome sequences, the genetic differences that have driven these changes in phenotype are only beginning to be understood. In this study, we use transposon-directed insertion-site sequencing to probe differences in gene requirements for competitive growth in rich media between these two closely related serovars. We identify a conserved core of 281 genes that are required for growth in both serovars, 228 of which are essential in *Escherichia coli*. We are able to identify active prophage elements through the requirement for their repressors. We also find distinct differences in requirements for genes involved in cell surface structure biogenesis and iron utilization. Finally, we demonstrate that transposon-directed insertion-site sequencing is not only applicable to the protein-coding content of the cell but also has sufficient resolution to generate hypotheses regarding the functions of non-coding RNAs (ncRNAs) as well. We are able to assign probable functions to a number of *cis*-regulatory ncRNA elements, as well as to infer likely differences in *trans*-acting ncRNA regulatory networks.**

INTRODUCTION

Salmonella enterica subspecies *enterica* serovars Typhi (*S.* Typhi) and Typhimurium (*S.* Typhimurium) are important human pathogens with distinctly different lifestyles. *S.* Typhi is host-restricted to humans and causes typhoid fever. This potentially fatal systemic illness affects at least 21 million people annually, primarily in developing countries (1–3) and is capable of colonizing the gall bladder creating asymptomatic carriers; such individuals are the primary source of this human restricted infection, exemplified by the case of 'Typhoid Mary' (4). *S.* Typhimurium, conversely, is a generalist, infecting a wide range of mammals and birds in addition to being a leading cause of foodborne gastroenteritis in human populations. Control of *S.* Typhimurium infection in livestock destined for the human food chain is of great economic importance, particularly in swine and cattle (5,6). Additionally, *S.* Typhimurium causes an invasive disease in mice, which has been used extensively as a model for pathogenicity in general and human typhoid fever specifically (7).

Despite this long history of investigation, the genomic factors that contribute to these differences in lifestyle remain unclear. More than 85% of predicted coding sequences are conserved between the two serovars in sequenced genomes of multiple strains (8–11). The horizontal acquisition of both plasmids and pathogenicity islands during the evolution of the salmonellae is believed to have impacted upon their disease potential. A 100 kb plasmid, encoding the *spv* (*Salmonella* plasmid

*To whom correspondence should be addressed. Tel: +44 1223 49 4448; Fax: +44 1223 49 4468; Email: lb14@sanger.ac.uk
Correspondence may also be addressed to John Wain. Tel: +44 1603 59 7567; Email: J.wain@uea.ac.uk
Present addresses:
Daniel J. Turner, Oxford Nanopore Technologies, Oxford, OX4 4GA, UK.
Minh-Duy Phan, Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, QLD 4072, Australia.
A. Keith Turner, Discuva Ltd., 260 Cambridge Science Park, Milton Road, Cambridge, CB4 0WE, UK.
John Wain, Norwich Medical School, University of East Anglia, NRP Innovation Centre, Norwich Research Park, Colney Lane, Norwich, Norfolk, NR4 7GJ, UK.
Paul P. Gardner, School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch, New Zealand.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

virulence) genes, is found in some *S.* Typhimurium strains and contributes significantly towards systemic infection in animal models (12,13). *S.* Typhi is known to have harboured IncHI1 plasmids conferring antibiotic resistance since the 1970's (14), and there is evidence that these strains present a higher bacterial load in the blood during human infection (15). Similar plasmids have been isolated from *S.* Typhimurium (16–18). *Salmonella* pathogenicity islands (SPI)-1 and -2 are common to both serovars and are required for invasion of epithelial cells [reviewed in (19)] and survival inside macrophages respectively (20,21). *S.* Typhi additionally incorporates SPI-7 and SPI-10, which contain the Vi surface antigen and a number of other putative virulence factors (22–24).

Acquisition of virulence determinants is not the sole explanation for the differing disease phenotypes displayed in humans by *S.* Typhimurium and *S.* Typhi; genome degradation is an important feature of the *S.* Typhi genome, in common with other host-restricted serovars such as *S.* Paratyphi A (humans) and *S.* Gallinarum (chickens). In each of these serovars, pseudogenes account for 4–7% of the genome (9,25–27). Loss of function has occurred in a number of *S.* Typhi genes that have been shown to encode intestinal colonization and persistence determinants in *S.* Typhimurium (28). Numerous sugar transport and degradation pathways have also been interrupted (9) but remain intact in *S.* Typhimurium, potentially underlying the restricted host niche occupied by *S.* Typhi.

In addition to its history as a model organism for pathogenicity, *S.* Typhimurium has recently served as a model organism for the elucidation of non-coding RNA (ncRNA) function (29). These include *cis*-acting switches, such as RNA-based temperature and magnesium ion sensors (30,31), together with a host of predicted metabolite-sensing riboswitches. Additionally, a large number of *trans*-acting small RNAs (sRNAs) have been identified within the *S.* Typhimurium genome (32), some with known roles in virulence (33). These sRNAs generally control a regulon of mRNA transcripts through an antisense binding mechanism mediated by the protein Hfq in response to stress. The functions of these molecules have generally been explored in either *S.* Typhimurium or *Escherichia coli*, and it is unknown how stable these functions and regulons are over evolutionary time (34).

Transposon mutagenesis has previously been used to assess the requirement of particular genes for cellular viability. The advent of next-generation sequencing has allowed simultaneous identification of all transposon insertion sites within libraries of up to 1 million independent mutants (35–38), enabling us to answer the basic question of which genes are required for *in vitro* growth with extremely fine resolution. By using transposon mutant libraries of this density, which in *S.* Typhi represents on average >80 unique insertions per gene (35), shorter regions of the genome can be interrogated, including ncRNAs (38). In addition, once these libraries exist, they can be screened through various selective conditions to further reveal which functions are required for growth/survival.

Using Illumina-based transposon-directed insertion-site sequencing [TraDIS (35)] with large mutant libraries of both *S.* Typhimurium and *S.* Typhi, we investigated whether these *Salmonellae* require the same protein-coding and ncRNA gene sets for competitive growth under laboratory conditions, and whether there are differences that reflect intrinsic differences in the pathogenic niches these bacteria inhabit.

MATERIALS AND METHODS

Strains

S. Typhimurium strain SL3261 was used to generate the large transposon mutant library, and contains a deletion relative to the parent strain SL1344. The 2166 bp deletion ranges from 153 bp within *aroA* (normally 1284 bp) to the last 42 bp of *cmk*, forming two pseudogenes and deleting the intervening gene SL0916 completely. For comparison, we used our previously generated *S.* Typhi Ty2 transposon library (35).

Annotation

For *S.* Typhimurium strain SL3261, we used feature annotations drawn from the SL1344 genome (EMBL-Bank accession FQ312003.1), ignoring the deleted *aroA*, *ycaL* and *cmk* genes. We re-analysed our *S.* Typhi Ty2 transposon library with features drawn from an updated genome annotation (EMBL-Bank accession AE014613.1). We supplemented the EMBL-Bank annotations with ncRNA annotations drawn from Rfam 10.1 (39), Sittka *et al.* (40), Chinni *et al.* (41), Raghavan *et al.* (42) and Kröger *et al.* (32). Selected protein-coding gene annotations were supplemented using the HMMER webserver (43) and Pfam (44).

Creation of *S.* Typhimurium transposon mutant library

S. Typhimurium was mutagenized using a Tn5-derived transposon as described previously (35). Briefly, the transposon was combined with the EZ-Tn5 transposase (Epicenter, Madison, USA) and electroporated into *S.* Typhimurium. Transformants were selected by plating on LB agar containing 15 μg/ml kanamycin and harvested directly from the plates following overnight incubation. A typical electroporation experiment generated a batch of between 50 000 and 150 000 individual mutants. Ten batches were pooled together to create a mutant library comprising ~930 000 transposon mutants.

DNA manipulations and sequencing

Genomic DNA was extracted from the library pool samples using tip-100 g columns and the genomic DNA buffer set from Qiagen (Crawley, UK). DNA was prepared for nucleotide sequencing as described previously (35). Before sequencing, a 22-cycle PCR was performed as previously described (35). Sequencing took place on a single end Illumina flowcell using an Illumina GAII sequencer, for 36 cycles of sequencing, using a custom sequencing primer and 2× Hybridization Buffer

(35). The custom primer was designed such that the first 10 bp of each read was transposon sequence.

Sequence analysis

The Illumina FASTQ sequence files were parsed for 100% identity to the 5′ 10 bp of the transposon (TAAGAGACA G). Sequence reads, which matched, were stripped of the transposon tag and subsequently mapped to the *S*. Typhimurium SL1344 or *S*. Typhi Ty2 chromosomes using Maq version maq-0.6.8 (45). Approximately 12 million sequence reads were generated from the sequencing run, which used two lanes on the Illumina flowcell. Precise insertion sites were determined using the output from the Maq mapview command, which gives the first nucleotide position to which each read mapped. The number and frequency of insertions mapping to each nucleotide in the appropriate genome was then determined.

Statistical analysis of required genes

The number of insertion sites for any gene is dependent on its length; therefore, the values were made comparable by dividing the number of insertion sites by the gene length, giving an 'insertion index' for each gene. As before (35), the distribution of insertion indices was bimodal, corresponding to the required (mode at 0) and non-required models. We fitted gamma distributions for the two modes using the R MASS library (http://www.r-project.org). Log_2-likelihood ratios (LLR) were calculated between the required and non-required models, and we called a gene required if it had an LLR of less than −2, indicating it was at least four times more likely according to the required model than the non-required model. 'Non-required' genes were assigned for an LLR of >2. Genes falling between the two thresholds were considered 'ambiguous' for the purpose of this analysis. This procedure lead to genes being called as required in *S*. Typhimurium when their insertion index was <0.020 and ambiguous between 0.020 and 0.027. The equivalent cut-offs for the *S*. Typhi library are 0.0147 and 0.0186, respectively.

We calculated a *P*-value for the observed number of insertion sites per gene using a Poisson approximation with rate R = N/G where N is the number of unique insert sites (549 086) and G is the number of bases in the genome (4 878 012). The *P*-value for at least X consecutive bases without an insert site is $e^{(-RX)}$, giving a 5% cut-off at 27 bp and a 1% cut-off at 41 bp.

For every gene *g* with $n_{g,A}$ reads observed in *S*. Typhi and $n_{g,B}$ reads observed in *S*. Typhimurium, we calculated the $\log_2$ fold change ratio $S_{g,A,B} = log_2\ [(n_{g,A}+100)/(n_{g,B}+100)]$. The correction of 100 reads smoothes out the high scores for genes with very low numbers of observed reads. We fitted a normal model to the mode +/− 2 sample standard deviations of the distribution of $S_{A,B}$ and calculated *P*-values for each gene according to the fit. We considered genes with a *P*-value of ≤0.05 under the normal model to be uniquely required by one serovar.

RESULTS AND DISCUSSION

TraDIS assay of every *Salmonella* Typhimurium protein-coding gene

Approximately 930 000 mutants of *S*. Typhimurium were generated using a *Tn5*-derived transposon. In all, 549 086 unique insertion sites were recovered from the mutant library using short-read sequencing with transposon-specific primers. This is a substantially higher density than the 371 775 insertions recovered from *S*. Typhi previously (35). The *S*. Typhimurium library contains an average of one insertion every 9 bp or >100 unique inserts per gene (Figure 1). The large number of unique insertion sites allowed every gene to be assayed; assuming random insertion across the genome, a region of 41 bp without an insertion was statistically significant ($P < 0.01$). As previously noted in *S*. Typhi, the distribution of length-normalized insertions per gene is bimodal (see Supplementary Figure S1), with one mode at 0. We interpret genes falling in to the distribution around this mode as being required for competitive growth within a mixed population under laboratory conditions (hereafter 'required'). Of these, 57 contained no insertions whatsoever and were mostly involved in core cellular processes (see Table 1, Supplementary Data Set).

There was a bias in the frequency of transposon insertion towards the origin of replication. This likely occurred as the bacteria were in exponential growth phase

Figure 1. Genome-wide transposon mutagenesis of *S*. Typhimurium. Circular plot showing gene content, distribution of required genes and insertion density along the *S*. Typhimurium chromosome. The outer scale is marked in megabases. Circular tracks range from 1 (outer track) to 6 (inner track). Track 1, all forward-strand genes (colour-coded according to function: dark blue, pathogenicity/adaptation; black, energy metabolism; red, information transfer; dark green, membranes/surface structures; cyan, degradation of macromolecules; purple, degradation of small molecules; yellow, central/intermediary metabolism; light blue, regulators; pink, phage/IS elements; orange, conserved hypothetical; pale green, unknown function; brown, pseudogenes.); track 2, all reverse-strand genes (colour-coded as for forward-strand genes); track 3, *S*. Typhimurium required genes (red); track 4, 56 genes required by *S*. Typhimurium, but not by *S*. Typhi (dark blue, see also Table 1); track 5, transposon insertion density; track 6, GC bias [(G − C)/(G + C)]), khaki indicates values >1; purple <1.

Table 1. Core gene functions in *S.* Typhimurium

Biological process	Sub-process	Required genes	Non-required genes
Cell division		***ftsALKQWYZ***, *minE*, ***mukB***, ***SL2391***	***ftsHJNX****, *minCD*, *sdiA*, *cedA*, *sulA*
DNA replication	Polymerases I, II and III	***dnaENQX***, ***holAB***	***polAB***, *holC****DE***
	Supercoiling	***gyrAB***, ***parCE***	
	Primosome-associated	***dnaBCGT***, ***priA***, ***ssb***	***priB****C*, ***rep***
Transcription	RNA polymerase	***rpoABC***	
	Sigma, elongation, anti- and termination factors	***nusBG***, ***rpoDH***, ***rho***	***nusA***, ***rpoE****NS*
Translation	tRNA-synthetases	***alaS***, ***argS***, ***asnS***, ***aspS***, ***cysS***, ***glnS***, ***gltX***, ***glyQS***, ***hisS***, ***ileS***, ***leuS***, ***lysS***, ***metG***, ***pheST***, ***proS***, ***serS***, ***thrS***, ***tyrS***, ***valS***,	*trpS*, *trpS2*
	Ribosome components	***rplBCDEFJKLMNOPQRSTUVWXY***, ***rpmABCDHI***, ***rpsABCDEFGHIJKLMNPQST***	*rplAI*, *rpmEE2*, *rpmFGJJ2*, *rps****OR****U****V*
	Initiation, elongation and peptide chain release factors	***fusA***, ***infABC***, ***prfAB***, ***tsf***, ***yrdC***	*efp*, *prfCH*, *selB*, *tuf*
Biosynthetic pathways			
Peptidoglycan		***murABCDEFGI***,	*ddl*, *ddlA*
Fatty acids		***accABCD***, ***fabABDGHIZ***	–

Protein-coding genes providing fundamental biological functions in *S.* Typhimurium. Genes in bold are required in *S.* Typhi (LLR between required and non-required models less than −2; see 'Materials and Methods' section). Asterisk indicates genes ambiguous in *S.* Typhimurium, having a LLR between −2 and 2.

immediately before transformation with the transposon. In this phase of growth, multiple replication forks would have been initiated, meaning genes closer to the origin were in greater copy number and hence more likely to be a target for insertion. We also observed a bias for transposon insertions in A + T rich regions, as was previously observed in the construction of an *S.* Typhi mutant library (35). However, the insertion density achieved is sufficient to discriminate between required and non-required genes easily. As was first seen in *S.* Typhi (35), we observed transposon insertions into genes upstream of required genes in the same operon, suggesting that most insertions do not have polar effects leading to the inactivation of downstream genes.

Analysis of the *S.* Typhimurium mutant library allowed us to identify 353 coding sequences required for growth under laboratory conditions, and 4112 non-required coding sequences (see Supplementary Data Set). We were unable to assign 65 genes to either the required or non-required category. Sixty of these genes, which we will refer to as 'ambiguous', had LLRs between −2 and 2. The final five unassigned genes had lengths <60 bases, and they were removed from the analysis. All other genes contained enough insertions or were of sufficient length to generate credible LLR scores. Thus, every gene was assayed, and we were able to draw conclusions for 98.7% of the coding genome in a single sequencing run (Figure 1).

Cross-species comparison of genes required for growth

Gene essentiality has previously been assayed in *Salmonella* using insertion-duplication mutagenesis (46). Knuth *et al.* estimated 490 genes are essential to growth in clonal populations, though 36 of these have subsequently been successfully deleted (47). Although TraDIS assays gene requirements after a brief period of competitive growth on rich media, we identify a smaller required set than Knuth *et al.* of ∼350 genes in each serovar, closer to current estimates of ∼300 essential genes in *E. coli* (48).

To demonstrate that TraDIS does identify genes known to have strong effects on growth, as well as to test our predictive power for determining gene essentiality, we compared our required gene sets in *S.* Typhimurium and *S.* Typhi to essential genes determined by systematic single-gene knockouts in the *E. coli* K-12 Keio collection (48). We identified orthologous genes in the three data sets by best reciprocal FASTA hits exhibiting >30% sequence identity for the amino acid sequences. Required orthologous genes identified in this manner share a significantly higher average percentage sequence identity with their *E. coli* counterparts than expected for a random set of orthologues, at ∼94% identity as compared with ∼87% for all orthologous genes. In 100 000 randomly chosen gene sets of the same size as our required set, we did not find a single set where the average shared identity exceeded 90%, indicating that required genes identified by TraDIS are more highly conserved at the nucleotide level than other orthologous protein-coding sequences.

Baba *et al.* (48) have defined an essentiality score for each gene in *E. coli* based on evidence from four experimental techniques for determining gene essentiality: targeted knockouts using λ-red mediated homologous recombination (48), genetic footprinting (49,50), large-scale chromosomal deletions (51) and transposon mutagenesis (52). Scores range from −4 to 3, with negative scores indicating evidence for non-essentiality and positive scores indicating evidence for essentiality. Comparing the overlap between essential gene sets in *E. coli*, *S.* Typhi and *S.* Typhimurium, we find a set of 228 *E. coli* genes, which have a Keio essentiality score of at least 0.5 (i.e. there is evidence for gene essentiality; see Figure 2) that have TraDIS-predicted required

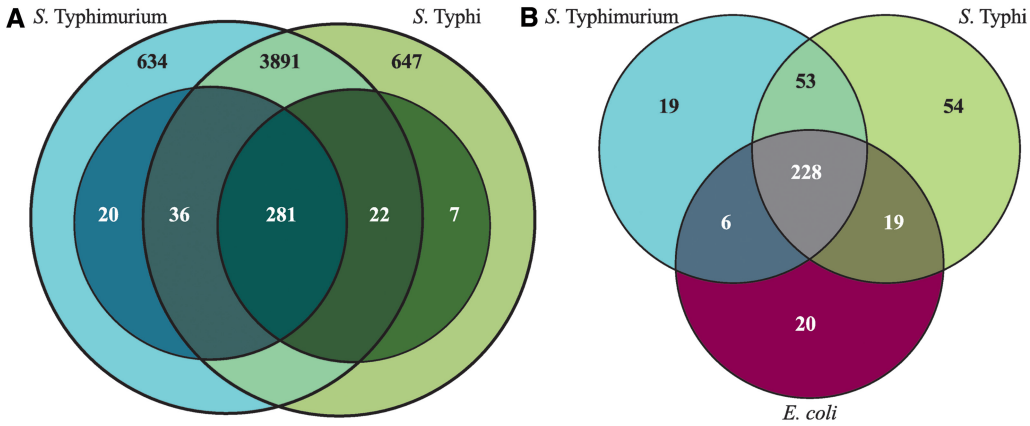

Figure 2. Comparison of required genes. Venn diagrams showing (**A**) the overlap of all genes (outer circles, light colours) and required genes (inner circles, dark colours) between *S.* Typhimurium and *S.* Typhi (excluding genes required in one serovar only, which do not have significantly different read-counts). Black numbers refer to all genes, white numbers to required genes. (**B**) the overlap of all required genes between *S.* Typhimurium (blue), *S.* Typhi (green) and *E. coli* (purple). White numbers refer to genes with Keio essentiality scores ≥0.5.

orthologues in both *S.* Typhi and *S.* Typhimurium, constituting ~85% of *E. coli* genes with evidence for essentiality indicating that gene requirements are largely conserved between these genera. Including orthologous genes that are only predicted to be essential by TraDIS in *S.* Typhi or *S.* Typhimurium raises this figure to nearly 93%. The majority of shared required genes between all three bacteria are responsible for fundamental cell processes, including cell division, transcription and translation. A number of key metabolic pathways are also represented, such as fatty acid and peptidoglycan biosynthesis (Table 1). A recent study in the alphaproteobacteria *Caulobacter crescentus* reported 210 shared essential genes with *E. coli*, despite *C. crescentus* sharing less than a third as many orthologous genes with *E. coli* as *Salmonella* serovars (38). This suggests the existence of a shared core of ~200 essential proteobacterial genes, with the comparatively rapid turnover of 150–250 'non-core' lineage-specific essential genes.

If we make the simplistic assumption that gene essentiality should be conserved between *E. coli* and *Salmonella*, we can use the overlap of our predictions with the Keio essential genes to provide an estimate of our TraDIS libraries' accuracy for predicting that a gene will be required in a clonal population. Of the 2632 orthologous *E. coli* genes, which have a Keio essentiality score of less than −0.5 (i.e. there is evidence for gene non-essentiality), only 33 are predicted to be required by TraDIS in both *Salmonella* serovars. *S.* Typhi contains the largest number of genes predicted by TraDIS to be required with *E. coli* orthologues with negative Keio essentiality scores. However, even if we assume these are all incorrect predictions of gene essentiality, this still gives a gene-wise false positive rate (FPR) of ~2.7% (81 of 2981 orthologues) and a positive predictive value (PPV) of ~75% (247 with essentiality scores ≥0.5 of 328 predictions with some Keio essentiality score). Under these same criteria, the *S.* Typhimurium data set has a lower gene-wise FPR of ~1.6% (51 of 3122 orthologues) and a higher PPV of ~82% (234 of 285 predictions as before), as we would expect given the library's higher insertion density. In reality, these FPRs and PPVs are only estimates; genes that are not essential in *E. coli* may become essential in the different genomic context of *Salmonella* serovars and vice versa, particularly in the case of *S.* Typhi where wide-spread pseudogene formation has eliminated potentially redundant pathways (26,27). Additionally, TraDIS will naturally over-predict essentiality in comparison with targeted knockouts, as our library creation protocol necessarily contains a short period of competitive growth between mutants during the recovery from electrotransformation and selection. As a consequence, genes that cause major growth defects, but not necessarily a complete lack of viability in clonal populations, may be reported as 'required'.

Serovar-specific genes required for growth

Many of the required genes present in only one serovar encoded phage repressors, for instance the cI proteins of Fels-2/SopE and ST35 (see Supplementary Tables S2 and S3). Repressors maintain the lysogenic state of prophage, preventing transcription of early lytic genes (53). Transposon insertions into these genes will relieve this repression and trigger the lytic cycle, resulting in cell death, and consequently mutants are not represented in the sequenced library. This again broadens the definition of 'required' genes; such repressors may not be required for cellular viability in the traditional sense, but once present in these particular genomes, their maintenance is required for continued viability as long as the rest of the phage remains intact.

Serovars Typhimurium and Typhi both contain eight apparent large phage-derived genomic regions (54,55). We were able to identify required repressors in all the intact lambdoid, P2-like and P22-like prophage in both genomes, including Gifsy-1, Gifsy-2 and Fels-2/SopE (see Supplementary Tables S2 and S3). With the exception of the SLP203 P22-like prophage in *S.* Typhimurium, all of these repressors lack the peptidase domain of the classical lambda repressor gene cI. This implies that the default anti-repression mechanism of *Salmonella* prophage may be more similar to a *trans*-acting

mechanism recently discovered in Gifsy phage (56) than to the phage lambda repressor's RecA-induced self-cleavage mechanism. We are also able to confirm that most phage remnants and fusions contained no active repressors, with the exception of the SLP281 degenerate P2-like prophage in *S*. Typhimurium. This degenerate prophage contains both intact replication and integration genes, but appears to lack tail and head proteins, suggesting it may depend on another phage for production of viral particles. Both genomes also encode P4-like satellite prophage, which rely on 'helper' phage for lytic functions and use a complex antisense-RNA based regulation mechanism for decision pathways regarding cell fate (57) using structural homologs of the IsrK (58) and C4 ncRNAs (59), known as seqA and CI RNA in the P4 literature, respectively. Although the mechanism of P4 lysogenic maintenance is not known, the IsrK-like ncRNAs of two potentially active P4-like prophage in *S*. Typhi are required under TraDIS. This sequence element has previously been shown to be essential for the establishment of the P4 lysogenic state (60), and we predict based on our observations that it may be necessary for lysogenic maintenance as well. The fact that some lambdoid prophage in *S*. Typhimurium encode non-coding genes structurally similar to the IsrK-C4 immunity system of P4 raises the possibility that these systems may be acting as a defense mechanism of sorts, protecting the prophage from predatory satellite phage capable of co-opting its lytic genes.

In addition to repressors, 4 prophage cargo genes in *S*. Typhimurium and one in *S*. Typhi are required (See Tables 2 and 3; Supplementary Tables S2 and S3). The *S*. Typhimurium prophage cargo genes encode a PhoPQ-regulated protein, a protein predicted to be involved in natural transformation, an endodeoxyribonuclease and a hypothetical protein. The *S*. Typhi prophage cargo gene encodes a protein containing the DNA-binding HIRAN domain (62) believed to be involved in the repair of damaged DNA. These warrant further investigation, as they are genes that have been recently acquired and become necessary for survival in rich media.

To compare differences between requirements for orthologous genes in both serovars, we calculated log-fold read ratios to eliminate genes, which were classified differently in *S*. Typhi and *S*. Typhimurium but did not have significantly different read densities (see 'Materials and Methods' section.) Even after this correction, 36 *S*. Typhimurium genes had a significantly lower frequency of transposon insertion compared with the equivalent genes in *S*. Typhi ($P < 0.05$), including four encoding hypothetical proteins (Table 2). This indicates that these gene products play a vital role in *S*. Typhimurium, but not in *S*. Typhi when grown under laboratory conditions.

A major difference between the two serovars is in the requirement for genes involved in cell wall biosynthesis (see Figure 3). A set of four genes (SL0702, SL0703, SL0706 and SL0707) in an operonic structure putatively involved in cell wall biogenesis is required in *S*. Typhimurium, but not in *S*. Typhi. The protein encoded by SL0706 is a pseudogene in *S*. Typhi (Ty2 unique ID: t2152) owing to a 1 bp deletion at codon 62 that causes a frameshift (Figure 4a). This operon contains an additional two pseudogenes in *S*. Typhi (t2154 and t2150), as well as a single different pseudogene (SL0700) in *S*. Typhimurium, indicating that this difference in gene requirements reflects the evolutionary adaptation of these serovars to their respective niches. Similarly, four genes (*rfbV*, *rfbX*, *rfbJ* and *rfbF*) within an O-antigen biosynthetic operon are required by *S*. Typhimurium, but not *S*. Typhi. There appears to have been a shuffling of O-antigen biosynthetic genes since the divergence between the two serovars, and *rfbJ*, encoding a CDP-abequose synthase, has been lost from *S*. Typhi altogether. These broader requirements for cell wall-associated biosynthetic and transporter genes suggest that surface structure biogenesis is of greater importance in *S*. Typhimurium.

We also identified seven genes from the shared pathogenicity island SPI-2 that appear to contain few or no transposon insertions only in *S*. Typhimurium under laboratory conditions. These genes (*spiC*, *sseA* and *ssaHIJT*) are thought to encode components of the SPI-2 type III secretion system apparatus (T3SS) (63). In addition, the effector genes *sseJ* and *sifB*, whose products are secreted through the SPI-2-encoded T3SS (64,65), also fell into the 'required' category in *S*. Typhimurium alone. All of these genes display high A + T nucleotide sequence and have been previously shown (in *S*. Typhimurium) to be strongly bound by the nucleoid-associated protein H-NS, encoded by *hns* (61,66). Therefore, rather than being 'required', it is instead possible that access for the transposon was sufficiently restricted that very few insertions occurred at these sites. In further support of this hypothesis, a comparison of the binding pattern of H-NS detected in studies using *S*. Typhimurium LT2 with the TraDIS results from the SPI-2 locus indicated that high regions of H-NS enrichment correlated well with both the *ssa* genes described here and with *sseJ* (61,66) (see Supplementary Figure S1). An earlier study also suggests that high-density DNA-binding proteins can block Mu, Tn*5* and Tn*10* insertion (67); however, a genome-wide study of the effects of H-NS binding on transposition would be necessary to confirm this effect.

Indeed, the generation of null *S*. Typhimurium mutants in *sseJ* and *sifB*, as well as many others generated at the SPI-2 locus suggest that these genes are not truly a requirement for growth in this serovar (65,68–70). Although this is a reminder that the interpretation of gene requirement needs to be made with care, the effect of H-NS on transposon insertion is not genome-wide. If this were the case, there would be an under-representation of transposon mutants in high A + T regions (known for H-NS binding), which is not what we observed. In total, only 21 required genes fall into the '*hns*-repressed' category described in Navarre *et al*. (61) (see Table 2; Supplementary Table S1); the remainder (almost 400) contained sufficient transposon insertions to conclude they were non-required. In addition, all SPI-1 genes that encode another T3SS and are of high A + T content were also found to be non-required. This phenomenon was not observed in *S*. Typhi, possibly because the strain used harbours the pHCM1 plasmid, which encodes the

Table 2. Genes uniquely required in *S.* Typhimurium

Ty inserts	Ty reads	SL inserts	SL reads	SL ID	SL gene length	Ty ID	Ty gene length	Name	Function
No orthologue in *S.* Typhi									
		18	123	SL0742	1269				putative cation transporter
		9	80	SL0830	516				conserved hypothetical protein
		4	21	SL0831	855				putative electron transfer flavoprotein (beta subunit)
		0	0	SL0950	323				predicted bacteriophage protein, potential phage repressor Gifsy-2
		11	75	SL1179	789			envF	lipoprotein
		3	18	SL1480	249				antitoxin Phd_YefM, type II toxin-antitoxin system
		4	32	SL1527	264			ydcX	putative inner membrane protein
		1	3	SL1560	717				putative membrane protein
		7	50	SL1601	859				putative transcriptional regulator (pseudogene)
		4	36	SL1799	201				bacteriophage encoded pagK (phoPQ-activated protein)
		5	22	SL1830A	434				conserved hypothetical protein (pseudogene)
		3	27	SL1967	677				predicted bacteriophage protein, potential phage repressor SLP203
		1	15	SL2045A	63			yoeI	short ORF
		17	107	SL2066	900			rfbJ	CDP-abequose synthase
		3	34	SL2549	209				endodeoxyribonuclease
		4	149	SL2593	449				putative DNA-binding protein, potential phage repressor Gifsy-1 SLP272
		3	7	SL2633	846				putative repressor protein, phage SLP281
			21	SL2695	978			smf	putative competence protein
		5	39	SL4132	291				hypothetical protein
		5	45	SL4354A	303				conserved hypothetical protein
Present in *S.* Typhi but required only in *S.* Typhimurium*									
36	474	5	26	SL0032	441	t0033	306		putative transcriptional regulator
71	349	11	48	SL0623	642	t2232	576	lipB	lipoate-protein ligase B
151	3546	10	64	SL0702	897	t2156	894		putative glycosyl transferase
194	3007	9	61	SL0703	1134	t2155	1134		galactosyltransferase
231	3499	15	67	SL0706	1779	t2152	1780		putative glycosyltransferase, cell wall biogenesis
84	1041	2	4	SL0707	834	t2151	834		putative glycosyltransferase, cell wall biogenesis
49	367	14	70	SL0722	1569	t2136	1569	cydA	cytochrome d ubiquinol oxidase subunit I
74	1613	5	22	SL1069	693	t1789	693		putative secreted protein
20	199	1	1	SL1203	150	t1146	156		hypothetical protein
20	290	1	5	SL1264	315	t1209	315		putative membrane protein
84	384	6	26	SL1327	402	t1261	384	spiC	putative pathogenicity island 2 secreted effector protein
66	769	5	35	SL1331	270	t1265	327	sseA	T3SS chaperone
36	307	2	5	SL1341	228	t1275	228	ssaH	putative pathogenicity island protein
47	407	1	3	SL1342	249	t1276	249	ssaI	putative pathogenicity island protein
144	3197	5	14	SL1343	750	t1277	750	ssaJ	putative pathogenicity island lipoprotein
63	847	5	26	SL1354	267	t1288	267	ssaS	putative type III secretion protein
73	762	4	44	SL1355	780	t1289	780	ssaT	putative type III secretion protein
30	226	12	48	SL1386	693	t1322	693	rnfE/ydgQ	Electron transport complex protein rnfE
265	3337	29	165	SL1473	1557	t1463	1557	pqaA	PhoPQ-activated protein
85	765	6	35	SL1532	951	t1511	951	sifB	putative virulence effector protein
22	156	16	174	SL1561	1227	t1534[a]	141	sseJ	Salmonella translocated effector protein (SseJ)
119	1639	10	44	SL1563	762	t1536	762		putative periplasmic amino acid-binding protein
107	2440	5	44	SL1564	648	t1537	648		putative ABC amino acid transporter permease
183	1646	20	118	SL1628	1355	t1612	1364		hypothetical protein
23	177	1	5	SL1659	183	t1640	183	yciG	conserved hypothetical protein
78	617	16	104	SL1684	1014	t1664	1014	Hnr	putative regulatory protein
37	277	4	25	SL1785	396	t1022	396		conserved hypothetical protein
166	2823	9	27	SL1793	915	t1016	915	pagO	inner membrane protein
28	311	3	22	SL1794	159	t1015	159		putative inner membrane protein
23	155	1	4	SL1823	972	t0988	972	msbB	lipid A acyltransferase
60	402	11	58	SL2064	1002	t0786	1002	rfbV	putative glycosyl transferase
87	524	7	59	SL2065	1293	t0785	1299	rfbX	putative O-antigen transporter
66	559	13	74	SL2069	774	t0780	774	rfbF	glucose-1-phosphate cytidylyltransferase
41	204	5	14	SL3828	1830	t3658	1830	glmS	glucosamine-fructose-6-phosphate aminotransferase
27	288	5	23	SL4250	288	t4220	288		putative GerE family regulatory protein
148	2633	16	89	SL4251	876	t4221	876		araC family regulatory protein

Genes determined to be uniquely required in *S.* Typhimurium. SL, *S.* Typhimurium; Ty, *S.* Typhi; inserts refer to the number of unique insertion sites within a gene; reads refer to the number of sequence reads over all insertions sites within a gene. Shaded rows indicate genes shown to be H-NS repressed in (61).

*P-value (associated with $\log_2$ read ratio) < 0.05.

[a]*sseJ* is a pseudogene in *S.* Typhi.

Table 3. Genes uniquely required in *S.* Typhi

SL inserts	SL reads	Ty inserts	Ty reads	Ty ID	Ty gene length	SL ID	SL gene length	Name	Function
No orthologue in *S.* Tm									
		1	5	t1332	132			malY	pseudogene
		2	32	t1920	405				possible repressor protein, prophage 10/Gifsy-2
		2	12	t3157	165				conserved hypothetical protein
		2	12	t3166	228				spurious ORF annotation overlapping the RnaseP/M1 RNA
		6	196	t3402	570			cI	repressor protein, cs 73 prophage
		4	58	t3415	741				HIRAN-domain family gene, potential DNA repair
		1	6	t4531	150				hypothetical secreted protein
Present in *S.* Typhimurium but required only in *S.* Typhi*									
199	1792	18	59	t0095	1287	SL0093	1287	surA	survival protein SurA precursor
45	498	3	22	t0123	459	SL0119	459	yabB/mraZ	conserved hypothetical protein
120	589	11	32	t0203	1281	SL0203	1281	hemL	glutamate-1-semialdehyde 2,1-aminomutase
123	982	2	25	t0224	1353	SL0224	1353	yaeL/rseP	Zinc metallopeptidase
67	452	1	14	t0270	576	SL2604	576	rpoE	RNA polymerase sigma-E factor
140	760	0	0	t0587	2286	SL2246	2286	nrdA	ribonucleoside-diphosphate reductase 1 alpha chain
113	641	15	42	t2140	2802	SL0718	2802	sucA	2-oxoglutarate dehydrogenase E1 component
116	753	13	36	t2177	1641	SL0680	1641	pgm	phosphoglucomutase
80	542	9	15	t2276	1008	SL0580	1008	fepD	ferric enterobactin transport protein FepD
93	591	2	2	t2277	990	SL0579	990	fepG	ferric enterobactin transport protein FepG
64	508	5	6	t2278	795	SL0578	795	fepC	ferric enterobactin transport ATP-binding protein FepC
201	1129	12	116	t2410	2355	SL0444	2355	lon	Lon protease
95	518	8	20	t2730	1062	SL2809	1062	recA[a]	recA protein
135	719	16	39	t2996	1992	SL3052	1947	tktA	transketolase
76	358	3	9	t3120	1434	SL3173	1434	rfaE	ADP-heptose synthase
213	1976	6	50	t3265	1071	SL3321	1071	degS	serine protease
43	448	3	10	t3326	606	SL3925	606	yigP	conserved hypothetical protein
124	571	17	36	t3384	2025	SL3872	2025	rep	ATP-dependent DNA helicase
175	1208	6	21	t3621	2787	SL3947	2787	polA	DNA polymerase I
117	797	9	13	t3808	1047	SL3677	1047	waaF	ADP-heptose-LPS heptosyltransferase II
176	1628	14	59	t4153	1080	SL4183	1080	alr	alanine racemase
140	1127	10	38	t4411	951	SL4294	951	miaA	tRNA delta-2-isopentenylpyrophosphate transferase

Genes determined to be uniquely required in *S.* Typhi. SL, *S.* Typhimurium; Ty, *S.* Typhi; inserts refer to the number of unique insertion sites within a gene; reads refer to the number of sequence reads over all insertions sites within a gene.
*P-value (associated with $\log_2$ read ratio) < 0.05.
[a]The assignment of *recA* as a required gene has been described previously (35), but briefly is believed to be due to the presence of the *priC* pseudogene in Typhi.

H-NS-like protein *sfh* and has been shown to affect H-NS binding (71,72).

Twenty-two *S.* Typhi genes had a significantly lower frequency of transposon insertion compared with orthologues in *S.* Typhimurium ($P < 0.05$), indicating that they are required only in *S.* Typhi for growth under laboratory conditions (Table 3), including the *fepBDGC* operon. This indicates a requirement for ferric [Fe(III)] rather than ferrous [Fe(II)] iron. This can be explained by the presence of Fe(III) in the bloodstream, where *S.* Typhi can be found during typhoid fever (15). These genes function to recover the ferric chelator enterobactin from the periplasm, acting with a number of proteins known to aid the passage of this siderophore through the outer membrane (73). It has long been noted that *aroA* mutants of *S.* Typhi, deficient in their ability to synthesize enterobactin, exhibit severe growth defects on complex media, whereas similar mutants of *S.* Typhimurium grow normally under the same conditions (74), though the mechanism has not been clear. Our results suggest that this difference in growth of *aroA* mutants is caused by a requirement for iron uptake through the *fep* system in *S.* Typhi. During host adaptation, *S.* Typhi has accumulated pseudogenes in many iron transport and response systems (27), presumably because they are not necessary for survival in the niche *S.* Typhi occupies in the human host, which may have led to this dependence on *fep* genes. In contrast, *S.* Typhimurium generally causes intestinal rather than systemic infection and is able to use a wider range of iron sources, including Fe(II), a soluble form of iron present under anaerobic conditions such as those found in the intestine (75).

TraDIS provides resolution sufficient to evaluate ncRNA contributions to fitness

Under a Poisson approximation to the transposon insertion process, a region of 41 (in *S.* Typhimurium) or 60

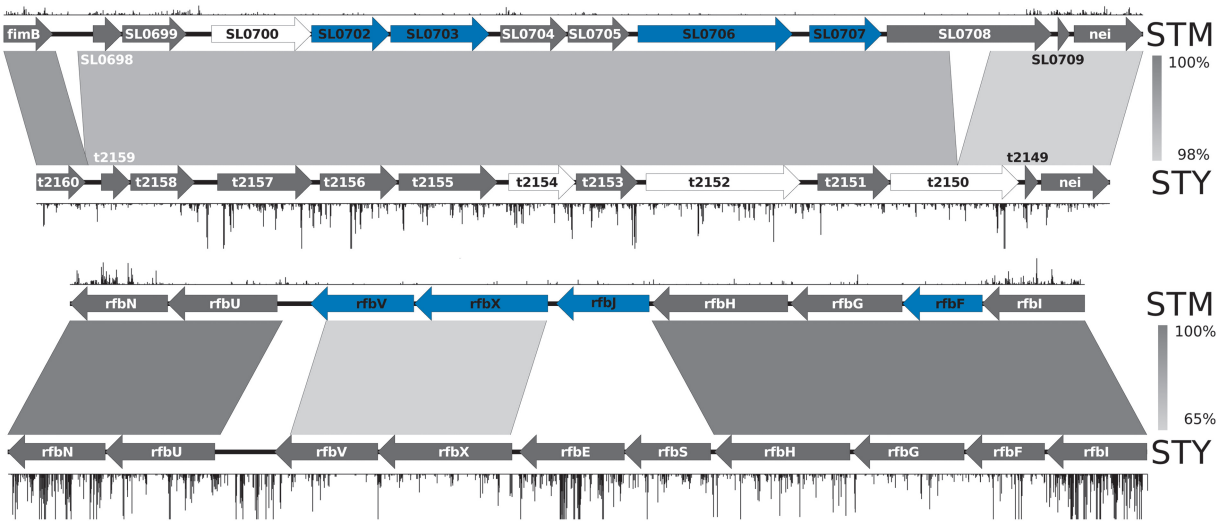

Figure 3. Comparison of cell surface operon structure and requirements. Diagram illustrating cell surface operons with different requirement patterns in *S*. Typhimurium and *S*. Typhi. The top figure is of an uncharacterized operon putatively involved in cell wall biogenesis, whereas the bottom figure shows a portion of the rfb operon involved in O-antigen biosynthesis. Plots along the top and bottom of each figure show insertions in *S*. Typhimurium and *S*. Typhi, respectively, with read depth on the *y*-axis with a maximum cut-off of 100 reads. Genes in blue are required in *S*. Typhimurium, genes in white are pseudogenes; others are in grey. Grey rectangles represent BLAST hits between orthologous genes, with percentage nucleotide identity coloured on the scale to the right of each figure.

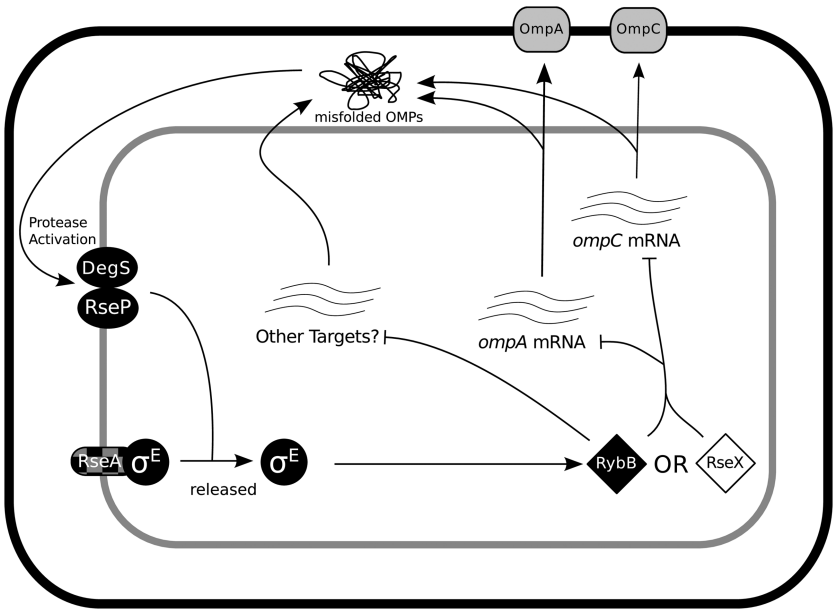

Figure 4. Proposed differences in sRNA utilization. Diagram illustrating inferred required sRNA regulatory networks under TraDIS. Molecules required in *S*. Typhi are in black and in *S*. Typhimurium are in white. RseA, in black/grey check, is ambiguous in *S*. Typhi. Non-required molecules are in grey. Diamonds indicate sRNAs, circles regulatory proteins, ovals proteases, oblong shapes are membrane-anchored proteins, and rounded squares are outer membrane porins.

bases (in *S*. Typhi) has only a 1% probability of not containing an insertion by chance. NcRNAs tend to be considerably shorter than their protein-coding counterparts, but this gives us sufficient resolution to assay most of the non-coding complement of the *Salmonella* genome. As a proof of principle, we performed an analysis of the best-understood class of small ncRNAs, the tRNAs. Francis Crick hypothesized that a single tRNA could recognize more than one codon through wobble recognition (76), where a non-canonical G-U base pair is formed between the first (wobble) position of the anticodon and the third nucleotide in the codon. As a result, some codons are covered by multiple tRNAs, whereas others are covered non-redundantly by a single tRNA. We expect that singleton wobble-capable tRNAs, i.e. wobble tRNAs which recognize a codon uniquely, will be required. In addition, we inferred the requirement for other tRNAs through the non-redundant coverage of

their codons and used this to benchmark our ability to use TraDIS to reliably interrogate short genomic intervals.

The *S*. Typhi and *S*. Typhimurium genomes encode 78 and 85 (plus one pseudogene) tRNAs, respectively, with 40 anticodons, as identified by tRNAscan-SE (77). In *S*. Typhi, 10 of 11 singleton wobble tRNAs are predicted to be required or ambiguous, compared with 16 tRNAs below the ambiguous LLR cut-off overall (significant enrichment at the 0.05 level, two-tailed Fisher's exact test *P*-value: 6.4e-08). Similarly in *S*. Typhimurium, 9 of 11 singleton wobble tRNAs are required or ambiguous compared with 15 required or ambiguous tRNAs overall, again showing a significant enrichment of required tRNAs in this subset (Fisher's exact test *P*-value: 5.2e-07). The one singleton wobble tRNA, which is consistently not required in both serovars is the tRNA-Pro(GGG), which occurs within a four-member codon family. It has previously been shown in *S*. Typhimurium that tRNA-Pro(UGG) can read all four proline codons *in vivo* owing to a cmo^5U34 modification to the anticodon, obviating the need for a functional tRNA-Pro(GGG) (78) and making this tRNA non-required. The other non-required singleton wobble tRNA in *S*. Typhimurium, tRNA-Leu(GAG), is similarly a member of a four-member codon family. We predict tRNA-Leu(TAG) may also be capable of recognizing all four leucine codons in this serovar; such a leucine 'four-way wobble' has been previously inferred in at least one other bacterial species (79,80).

Of the six required non-wobble tRNAs in each serovar, four are shared. These include two non-wobble singleton tRNAs covering codons uniquely, as well as a tRNA with the ATG anticodon, which is post-transcriptionally modified by the required protein mesJ/tilS to recognize the isoleucine codon ATA (80). An additional two required tRNAs in both serovars, one shared and one with a differing anticodon, contain Gln anticodons and are part of a polycistronic tRNA operon containing other required tRNAs. This operon is conserved in *E. coli* with the exception of an additional tRNA-Gln at the 3′ end that has been lost in the *Salmonella* lineage. It is possible that transposon insertions early in the operon may interfere with processing of the polycistronic transcript in to mature tRNAs. Finally, we do not observe insertions in a tRNA-Met and a tRNA-Val in *S*. Typhi and *S*. Typhimurium, respectively.

Using this analysis of the tRNAs, we estimate a worst-case PPV for these short molecules (~76 bases) at 81%, in line with our previous estimates for conserved protein-coding genes, and a FPR of <4%, higher than for protein-coding genes but still well within the typical tolerance of high-throughput experiments. This assumes that the 'required' operonic tRNA-Glns and the serovar-specific tRNA-Met and tRNA-Val are all false positives; it is not clear that this is in fact the case.

Surveying the shared required ncRNA content of both serovars (see Table 4), we find that the RNA components of the signal recognition particle (SRP) and RNaseP, two universally conserved ncRNAs, are required as expected. The SRP is an essential component of the cellular secretion machinery, whereas RnaseP is necessary for the maturation of tRNAs. We also find a number of required known and potential *cis*-regulatory molecules associated with genes required for growth under laboratory conditions in both serovars. The RFN riboswitch controls *ribB*, a 3,4-dihydroxy-2-butanone 4-phosphate synthase involved in riboflavin biosynthesis, in response to flavin mononucleotide concentrations (83). Additionally, we are able to assign putative functions to a number of previously uncharacterized required non-coding transcripts through their 5′ association with required genes. SroE, a 90 nt molecule discovered in an early sRNA screen (84), is consistently located at the 5′ end of the required *hisS* gene across its phylogenic distribution in the Enterobacteriaceae. Given this consistent association and the function of HisS as a histidyl-tRNA synthetase, we hypothesize that this region may act in a manner similar to a T-box leader, inducing or repressing expression in response to tRNA-His levels. The *thrU* leader sequence, recently discovered in a deep-sequencing screen of *E. coli* (42), appears to regulate a polycistronic operon of required singleton wobble tRNAs. Three additional required *cis*-regulatory elements, t44, S15 and StyR-8, are associated with required ribosomal proteins, highlighting the central role ncRNA elements play in regulating fundamental cellular processes.

The sRNAs required for competitive growth

Inferring functions for potential *trans*-acting ncRNA molecules, such as anti-sense binding sRNAs, from requirement patterns alone is more difficult than for *cis*-acting elements, as we cannot rely on adjacent genes to provide any information. It is also important to keep in mind that TraDIS assays requirements after a brief competition within a large library of mutants on permissive media. This may be particularly important when surveying the bacterial sRNAs, which are known to participate in responses to stress (29).

This is demonstrated by two sRNAs involved in the σ^E-mediated extracytoplamic stress response, RybB and RseX, both of which can be successfully knocked out in *S*. Typhimurium (101). In *S*. Typhi, *rpoE* is required, as it also is in *E. coli* (48,102). However, in *S*. Typhimurium, *rpoE* tolerates a heavy insertion load, implying that σ^E mutants are not disadvantaged in competitive growth. In *S*. Typhimurium, the sRNA RseX is required. Overexpression of RseX has previously been shown to compensate for σ^E essentiality in *E. coli* by degrading *ompA* and *ompC* transcripts (95). This suggests that RseX may also be short-circuiting the σ^E stress response network in *S*. Typhimurium (Figure 4). To our knowledge, this is the first evidence of a native (i.e. not experimentally induced) activity of RseX.

S. Typhi on the other hand requires σ^E along with its activating proteases RseP and DegS and anchoring protein RseA, as well as the σ^E-dependent sRNA RybB, which also regulates OmpA and OmpC in *S*. Typhimurium, along with a host of other OMPs (103). It is unclear why the σ^E response is required in *S*. Typhi, but not *S*. Typhimurium, though it may partially be due to the major differences in the cell wall and outer

Table 4. Candidate required ncRNAs greater than 60 nt in length, excluding rRNA and tRNA

Element name	Rfam accession	Function	Hfq-binding	Downstream protein-coding gene(s)	Downstream gene required	References
Required or ambiguous in both *S.* Typhi and *S.* Typhimurium						
SRP	RF00169	RNA component of the signal recognition particle				(81)
RNase P	RF00010	RNA component of RNase P		ybaZ	N	(82)
RFN	RF00050	FMN-sensing riboswitch controlling the ribB gene		ribB	Y	(83)
SroE	RF00371	Putative cis-regulatory element controlling the hisS gene		hisS	Y	(84)
ThrU Leader	NA	Putative cis-regulatory element controlling the ThrU tRNA operon				(42)
t44	RF00127	Cis-regulatory element controlling the ribosomal rpsB gene		rpsB	Y	(85-87)
S15[b]	RF00114	Translational regulator of the ribosomal S15 protein		rpsO	Y	(88)
StyR-8	NA	Putative cis-regulatory element controlling the ribosomal rpmB gene		rpmB	Y	(41)
MicA	RF00078	sRNA involved in cellular response to extracytoplasmic stress	Y	luxS	N	(29)
DsrA[b]	RF00014	sRNA regulator of H-NS	Y	mngB	N	(89)
STnc10	NA	Putative sRNA		nhaA	N	(40)
STnc60[b]	NA	Putative sRNA				(40)
STnc840	NA	Verified sRNA derived from 3' UTR of the flgL gene	Y			(90)
IS0420[a,b]	NA	Putative ncRNA		rmf	N	(42,91)
RGO0[b]	NA	Putative sRNA identified in E. coli				(42)
Required or ambiguous in *S.* Typhimurium only						
rne5	RF00040	RNase E autoregulatory 5' element		rne	Y	(92)
RydC	RF00505	sRNA regulator of the yejABEF ABC transporter	Y			(93)
RydB	RF00118	Putative ncRNA				(94)
STnc510	NA	Putative sRNA		pagD/pagC	Y/N	(40)
STnc460[b]	NA	Putative sRNA				(40)
STnc170	NA	Putative sRNA		SL1458	N	(40)
STnc130	NA	Putative sRNA		dmsA	N	(40)
RseX	RF01401	sRNA regulator of OmpA and OmpC	Y			(95)
IsrJ	RF01393	sRNA regulator of SPI-1 effector protein secretion				(40,58)
IsrI	RF01392	Island-encoded Hfq-binding sRNA	Y	SL1028	Y	(40,58,90)
Required or ambiguous in *S.* Typhi only						
RybB	RF00110	sRNA involved in cellular response to extracytoplasmic stress	Y			(29)
tk5[a]	NA	Putative ncRNA				(42,96)
STnc750	NA	Verified sRNA	Y	SpeB	N	(32,90)
StyR-44[a]	RF01830	Putative multicopy (2/6 copies required in *S.* Typhi) ncRNA associated with ribosomal RNA operon		23S rRNA	N	(41)
tp2	NA	Putative ncRNA		aceE	N	(42,96)
RdlD	RF01813	RdlD RNA anti-toxin, 1/2 copies required in *S.* Typhi				(97)
STnc120[a]	NA	Putative sRNA				(40)
tp28[a]	NA	Putative ncRNA		fur	N	(42,96)
Phe Leader[a]	RF01859	Phenylalanine peptide leader sequence associated with the required PheST operon		PheS	Y	(98)
RimP Leader	RF01770	Putative cis-regulator of the rimP-nusA-infB operon		rimP	Y	(99)
GlmY	RF00128	Trans-acting regulator of the GlmS gene				(100)

Known and putative non-coding elements classified as required or ambiguous in this screen. Required ncRNAs have an LLR between required and non-required models of less than −2; see 'Materials and Methods' section.
[a,b]ncRNAs that are amibiguous (LLR between −2 and 2) in *S.* Typhi([a]) or in *S.* Typhimurium([b]). Hfq-binding annotations are taken from (90). The downstream protein-coding genes columns report annotated CDS or ribosomal RNA start sites within 100 bases of each candidate required non-coding element on either strand and whether these downstream sequences are also classified as required.

membrane between the two serovars. In addition, there are significant differences in the OMP content of the *S*. Typhi and *S*. Typhimurium membranes that may be driving alternative mechanisms for coping with membrane stress. For instance, *S*. Typhi completely lacks OmpD, a major component of the *S*. Typhimurium outer membrane (104) and a known target of RybB (29).

Two additional sRNAs involved in stress response are also required by both *S*. Typhi and *S*. Typhimurium. The first, MicA, is known to regulate *ompA* and the *lamB* porin-coding gene in *S*. Typhimurium (105), contributing to the extracytoplasmic stress response. The second, DsrA, has been shown to negatively regulate the nucleoid-forming protein H-NS and enhance translation of the stationary-phase alternative sigma factor σ^S in *E. coli* (89), though its regulation of σ^S does not appear to be conserved in *S*. Typhimurium (106). Both have been previously deleted in *S*. Typhimurium and thus are not essential. H-NS knockouts have previously been shown to have severe growth defects in *S*. Typhimurium that can be rescued by compensatory mutations in either the *phoPQ* two-component system or *rpoS*, implying that the lack of H-NS is allowing normally silenced detrimental regions to be transcribed (61). As MicA has recently been shown to negatively regulate *phoPQ* expression in *E. coli* (107), it is tempting to speculate that MicA may be moderating the effects of DsrA-induced H-NS repression; however, it is currently unclear whether sRNA regulons are sufficiently conserved between *E. coli* and *S. enterica* to justify this hypothesis.

CONCLUSION

The extremely high resolution of TraDIS has allowed us to assay gene requirements in two very closely related *Salmonellae* with different host ranges. We found, under laboratory conditions, that 58 genes present in both serovars were required in only one, suggesting that identical gene products do not necessarily have the same phenotypic effects in the two different serovar backgrounds. Many of these genes occur in genomic regions or metabolic systems, which contain pseudogenes and/or have undergone reorganization since the divergence of *S*. Typhi and *S*. Typhimurium, demonstrating the complementarity of TraDIS and phylogenetic analysis. These changes may, in part, explain differences observed in the pathogenicity and host specificity of these two serovars. In particular, *S*. Typhimurium showed a requirement for cell surface structure biosynthesis genes; this may be partially explained by the fact that *S*. Typhi expresses the Vi-antigen, which masks the cell surface, though these genes are not required for survival in our assay. *S*. Typhi on the other hand has a requirement for iron uptake through the *fep* system, which enables ferric enterobactin transport. This dependence on enterobactin suggests that *S*. Typhi is highly adapted to the iron-scarce environments it encounters during systemic infections. Furthermore, this appears to represent a single point of failure in the *S*. Typhi iron utilization pathways and may present an attractive target for narrow-spectrum antibiotics.

Of the ~4500 protein-coding genes present in each serovar, only ~350 were sufficiently depleted in transposon insertions to be classified as required for growth in rich media. This means that >92% of the coding genome has sufficient insertion density to be queried in future assays. Dense transposon mutagenesis libraries have been used to assay gene requirements under conditions relevant for infection, including *S*. Typhi survival in bile (35), *Mycobacterium tuberculosis* catabolism of cholesterol (108), drug resistance in *Pseudomonas aeruginosa* (109) and *Haemophilus influenzae* survival in the lung (110). We expect that parallel experiments querying gene requirements under the same conditions in both serovars examined in this study will yield further insights in to the differences in the infective process between Typhi and Typhimurium and ultimately the pathways that underlie host-adaptation.

Both serovars possess substantial complements of horizontally acquired DNA. We have been able to use TraDIS to assay these recently acquired sequences. In particular, we have been able to identify, on a chromosome wide scale, active prophage through the requirement for their repressors. The P4 phage uses an RNA-based system to make decisions regarding cell fate, and structurally similar systems are used by P1, P7 and N15 phage (111,112). C4-like transcripts have been regarded as the primary repressor of lytic functions, though the IsrK-like sequence is known to be essential to the establishment of lysogeny in P4 and is transcribed in at least two phage types (60,112). Our observations in *S*. Typhi suggest an important role for the IsrK-like sequence in maintenance of the lysogenic state in P4-like phage, though the mechanism remains unclear.

Recent advances in high-throughput sequencing have greatly enhanced our ability to detect novel transcripts, such as ncRNAs and short open reading frames (sORFs). Our ability to identify these transcripts now far out-strips our ability to experimentally characterize these sequences. There have been previous efforts at high-throughput characterization of bacterial sRNAs and sORFs in enteric bacteria; however, these have relied on labour-intensive directed knockout libraries (47,113). Here, we have demonstrated that TraDIS has sufficient resolution to reliably query genomic regions as short as 60 bases, in agreement with a recent high-throughput transposon mutagenesis study in the alphaproteobacteria *Caulobacter crescentus* (38). Our method has the major advantage that library construction does not rely on genome annotation, and newly discovered elements can be surveyed with no further laboratory work.

We have been able to assign putative functions to a number of ncRNAs using TraDIS though consideration of their genomic and experimental context. In addition, ncRNA characterization generally is done in model organisms like *E. coli* or *S*. Typhimurium, and it is unclear how stable ncRNA regulatory networks are over evolutionary time. By assaying two serovars of *Salmonella* with the same method under the same conditions, we have seen hints that there may be differences in sRNA regulatory

networks between *S*. Typhi and *S*. Typhimurium. In particular, we have found that under the same experimental conditions, *S*. Typhi appears to rely on the σ^E stress response pathway, whereas *S*. Typhimurium does not; it is tempting to speculate that this difference in stress response is mediated by the observed difference in requirement for two sRNAs, RybB and RseX. We believe that this combination of high-throughput transposon mutagenesis with a careful consideration of the systems context of individual genes provides a powerful tool for the generation of functional hypotheses. We anticipate that the construction of TraDIS libraries in additional organisms, as well as the passing of these libraries through relevant experimental conditions, will provide further insights into the function of bacterial ncRNAs in addition to the protein-coding gene complement.

ACCESSION NUMBERS

ERA000097, ERA000217.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3, Supplementary Figures 1–2 and Supplementary Data Set.

ACKNOWLEDGEMENTS

The authors would like to thank Leopold Parts for assistance with the statistical analysis of required genes under laboratory conditions, Derek Pickard for discussions of phage biology and Amy Cain for comments on the manuscript. The *S*. Typhimurium nucleotide sequencing data have been deposited in the European Short Read Archive under accession number ERA000217. The *S*. Typhi data can be accessed at ERA000097.

FUNDING

Wellcome Trust [WT076964, WT079643 and WT098051]; Medical Research Council. Funding for open access charge: Wellcome Trust.

Conflict of interest statement. None declared.

REFERENCES

1. Crump,J.A., Luby,S.P. and Mintz,E.D. (2004) The global burden of typhoid fever. *Bull World Health Organ.*, **82**, 346–353.
2. Bhutta,Z.A. and Threlfall,J. (2009) Addressing the global disease burden of typhoid fever. *JAMA*, **302**, 898–899.
3. Kothari,A., Pruthi,A. and Chugh,T.D. (2008) The burden of enteric fever. *J. Infect. Dev. Ctries.*, **2**, 253–259.
4. Soper,G.A. (1939) The curious career of typhoid mary. *Bull N.Y. Acad. Med.*, **15**, 698–712.
5. CDC. (2009) *Salmonella* surveillance: Annual summary.
6. Majowicz,S.E., Musto,J., Scallan,E., Angulo,F.J., Kirk,M., O'Brien,S.J., Jones,T.F., Fazil,A. and Hoekstra,R.M. (2010) The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clin. Infect. Dis.*, **50**, 882–889.
7. Santos,R.L., Zhang,S., Tsolis,R.M., Kingsley,R.A., Adams,L.G. and Baumler,A.J. (2001) Animal models of *Salmonella* infections: enteritis versus typhoid fever. *Microbes Infect.*, **3**, 1335–1344.
8. McClelland,M., Sanderson,K.E., Spieth,J., Clifton,S.W., Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F. *et al.* (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*, **413**, 852–856.
9. Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D., Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T. *et al.* (2001) Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature*, **413**, 848–852.
10. Holt,K.E., Parkhill,J., Mazzoni,C.J., Roumagnac,P., Weill,F.X., Goodhead,I., Rance,R., Baker,S., Maskell,D.J., Wain,J. *et al.* (2008) High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat. Genet.*, **40**, 987–993.
11. Deng,W., Liou,S.R., Plunkett,G. 3rd, Mayhew,G.F., Rose,D.J., Burland,V., Kodoyianni,V., Schwartz,D.C. and Blattner,F.R. (2003) Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. *J. Bacteriol.*, **185**, 2330–2337.
12. Gulig,P.A. and Curtiss,R. 3rd (1987) Plasmid-associated virulence of *Salmonella* Typhimurium. *Infect. Immun.*, **55**, 2891–2901.
13. Gulig,P.A. and Doyle,T.J. (1993) The *Salmonella* Typhimurium virulence plasmid increases the growth rate of salmonellae in mice. *Infect. Immun.*, **61**, 504–511.
14. Phan,M.D., Kidgell,C., Nair,S., Holt,K.E., Turner,A.K., Hinds,J., Butcher,P., Cooke,F.J., Thomson,N.R., Titball,R. *et al.* (2009) Variation in *Salmonella enterica* serovar Typhi IncHI1 plasmids during the global spread of resistant typhoid fever. *Antimicrob. Agents Chemother.*, **53**, 716–727.
15. Wain,J., Diep,T.S., Ho,V.A., Walsh,A.M., Nguyen,T.T., Parry,C.M. and White,N.J. (1998) Quantitation of bacteria in blood of typhoid fever patients and relationship between counts and clinical features, transmissibility, and antibiotic resistance. *J. Clin. Microbiol.*, **36**, 1683–1687.
16. Datta,N. (1962) Transmissible drug resistance in an epidemic strain of *Salmonella* Typhimurium. *J. Hyg. (Lond.)*, **60**, 301–310.
17. Holt,K.E., Thomson,N.R., Wain,J., Phan,M.D., Nair,S., Hasan,R., Bhutta,Z.A., Quail,M.A., Norbertczak,H., Walker,D. *et al.* (2007) Multidrug-resistant *Salmonella enterica* serovar Paratyphi A harbors IncHI1 plasmids similar to those found in serovar Typhi. *J. Bacteriol.*, **189**, 4257–4264.
18. Cain,A.K. and Hall,R.M. (2012) Evolution of a multiple antibiotic resistance region in IncHI1 plasmids: reshaping resistance regions in situ. *J. Antimicrob. Chemother.*, **67**, 2848–2853.
19. Darwin,K.H. and Miller,V.L. (1999) Molecular basis of the interaction of *Salmonella* with the intestinal mucosa. *Clin. Microbiol. Rev.*, **12**, 405–428.
20. Ochman,H., Soncini,F.C., Solomon,F. and Groisman,E.A. (1996) Identification of a pathogenicity island required for *Salmonella* survival in host cells. *Proc. Natl Acad. Sci. USA*, **93**, 7800–7804.
21. Shea,J.E., Hensel,M., Gleeson,C. and Holden,D.W. (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella* Typhimurium. *Proc. Natl Acad. Sci. USA*, **93**, 2593–2597.
22. Pickard,D., Wain,J., Baker,S., Line,A., Chohan,S., Fookes,M., Barron,A., Gaora,P.O., Chabalgoity,J.A., Thanky,N. *et al.* (2003) Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding *Salmonella enterica* pathogenicity island SPI-7. *J. Bacteriol.*, **185**, 5055–5065.
23. Seth-Smith,H.M. (2008) SPI-7: *Salmonella*'s Vi-encoding Pathogenicity Island. *J. Infect. Dev. Ctries.*, **2**, 267–271.
24. Townsend,S.M., Kramer,N.E., Edwards,R., Baker,S., Hamlin,N., Simmonds,M., Stevens,K., Maloy,S., Parkhill,J., Dougan,G. *et al.* (2001) *Salmonella enterica* serovar Typhi possesses a unique repertoire of fimbrial gene sequences. *Infect. Immun.*, **69**, 2894–2901.
25. Thomson,N.R., Clayton,D.J., Windhorst,D., Vernikos,G., Davidson,S., Churcher,C., Quail,M.A., Stevens,M., Jones,M.A., Watson,M. *et al.* (2008) Comparative genome analysis of *Salmonella* Enteritidis PT4 and Salmonella Gallinarum 287/91

provides insights into evolutionary and host adaptation pathways. *Genome Res.*, **18**, 1624–1637.
26. Holt,K.E., Thomson,N.R., Wain,J., Langridge,G.C., Hasan,R., Bhutta,Z.A., Quail,M.A., Norbertczak,H., Walker,D., Simmonds,M. *et al.* (2009) Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi. *BMC Genomics*, **10**, 36.
27. McClelland,M., Sanderson,K.E., Clifton,S.W., Latreille,P., Porwollik,S., Sabo,A., Meyer,R., Bieri,T., Ozersky,P., McLellan,M. *et al.* (2004) Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. *Nat. Genet.*, **36**, 1268–1274.
28. Kingsley,R.A., Humphries,A.D., Weening,E.H., De Zoete,M.R., Winter,S., Papaconstantinopoulou,A., Dougan,G. and Baumler,A.J. (2003) Molecular and phenotypic analysis of the CS54 island of *Salmonella enterica* serotype Typhimurium: identification of intestinal colonization and persistence determinants. *Infect. Immun.*, **71**, 629–640.
29. Vogel,J. (2009) A rough guide to the non-coding RNA world of *Salmonella*. *Mol. Microbiol.*, **71**, 1–11.
30. Waldminghaus,T., Heidrich,N., Brantl,S. and Narberhaus,F. (2007) FourU: a novel type of RNA thermometer in *Salmonella*. *Mol. Microbiol.*, **65**, 413–424.
31. Cromie,M.J., Shi,Y., Latifi,T. and Groisman,E.A. (2006) An RNA sensor for intracellular Mg(2+). *Cell*, **125**, 71–84.
32. Kroger,C., Dillon,S.C., Cameron,A.D., Papenfort,K., Sivasankaran,S.K., Hokamp,K., Chao,Y., Sittka,A., Hebrard,M., Handler,K. *et al.* (2012) The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc. Natl Acad. Sci. USA*, **109**, E1277–E1286.
33. Hebrard,M., Kroger,C., Srikumar,S., Colgan,A., Handler,K. and Hinton,J. (2012) sRNAs and the virulence of *Salmonella enterica* serovar Typhimurium. *RNA Biol.*, **9**, 437–445.
34. Richter,A.S. and Backofen,R. (2012) Accessibility and conservation: General features of bacterial small RNA-mRNA interactions? *RNA Biol.*, **9**, 954–965.
35. Langridge,G.C., Phan,M.D., Turner,D.J., Perkins,T.T., Parts,L., Haase,J., Charles,I., Maskell,D.J., Peters,S.E., Dougan,G. *et al.* (2009) Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.*, **19**, 2308–2316.
36. Goodman,A.L., McNulty,N.P., Zhao,Y., Leip,D., Mitra,R.D., Lozupone,C.A., Knight,R. and Gordon,J.I. (2009) Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe.*, **6**, 279–289.
37. van Opijnen,T., Bodi,K.L. and Camilli,A. (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods.*, **6**, 767–772.
38. Christen,B., Abeliuk,E., Collier,J.M., Kalogeraki,V.S., Passarelli,B., Coller,J.A., Fero,M.J., McAdams,H.H. and Shapiro,L. (2011) The essential genome of a bacterium. *Mol. Syst. Biol.*, **7**, 528.
39. Burge,S.W., Daub,J., Eberhardt,R., Tate,J., Barquist,L., Nawrocki,E.P., Eddy,S.R., Gardner,P.P. and Bateman,A. (2013) Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.*, **41**, D226–D232.
40. Sittka,A., Sharma,C.M., Rolle,K. and Vogel,J. (2009) Deep sequencing of *Salmonella* RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes. *RNA Biol.*, **6**, 266–275.
41. Chinni,S.V., Raabe,C.A., Zakaria,R., Randau,G., Hoe,C.H., Zemann,A., Brosius,J., Tang,T.H. and Rozhdestvensky,T.S. (2010) Experimental identification and characterization of 97 novel npcRNA candidates in *Salmonella enterica* serovar Typhi. *Nucleic Acids Res.*, **38**, 5893–5908.
42. Raghavan,R., Groisman,E.A. and Ochman,H. (2011) Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.*, **21**, 1487–1497.
43. Finn,R.D., Clements,J. and Eddy,S.R. (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.*, **39**, W29–W37.
44. Punta,M., Coggill,P.C., Eberhardt,R.Y., Mistry,J., Tate,J., Boursnell,C., Pang,N., Forslund,K., Ceric,G., Clements,J. *et al.* (2012) The Pfam protein families database. *Nucleic Acids Res.*, **40**, D290–D301.
45. Li,H., Ruan,J. and Durbin,R. (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.*, **18**, 1851–1858.
46. Knuth,K., Niesalla,H., Hueck,C.J. and Fuchs,T.M. (2004) Large-scale identification of essential *Salmonella* genes by trapping lethal insertions. *Mol. Microbiol.*, **51**, 1729–1744.
47. Santiviago,C.A., Reynolds,M.M., Porwollik,S., Choi,S.H., Long,F., Andrews-Polymenis,H.L. and McClelland,M. (2009) Analysis of pools of targeted *Salmonella* deletion mutants identifies novel genes affecting fitness during competitive infection in mice. *PLoS Pathog.*, **5**, e1000477.
48. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
49. Gerdes,S.Y., Scholle,M.D., Campbell,J.W., Balazsi,G., Ravasz,E., Daugherty,M.D., Somera,A.L., Kyrpides,N.C., Anderson,I., Gelfand,M.S. *et al.* (2003) Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J. Bacteriol.*, **185**, 5673–5684.
50. Tong,X., Campbell,J.W., Balazsi,G., Kay,K.A., Wanner,B.L., Gerdes,S.Y. and Oltvai,Z.N. (2004) Genome-scale identification of conditionally essential genes in *E. coli by DNA microarrays*. *Biochem. Biophys. Res. Commun.*, **322**, 347–354.
51. Hashimoto,M., Ichimura,T., Mizoguchi,H., Tanaka,K., Fujimitsu,K., Keyamura,K., Ote,T., Yamakawa,T., Yamazaki,Y., Mori,H. *et al.* (2005) Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. *Mol. Microbiol.*, **55**, 137–149.
52. Kang,Y., Durfee,T., Glasner,J.D., Qiu,Y., Frisch,D., Winterberg,K.M. and Blattner,F.R. (2004) Systematic mutagenesis of the *Escherichia coli* genome. *J. Bacteriol.*, **186**, 4921–4930.
53. Echols,H. and Green,L. (1971) Establishment and maintenance of repression by bacteriophage lambda: the role of the cI, cII, and c3 proteins. *Proc. Natl Acad. Sci. USA*, **68**, 2190–2194.
54. Thomson,N., Baker,S., Pickard,D., Fookes,M., Anjum,M., Hamlin,N., Wain,J., House,D., Bhutta,Z., Chan,K. *et al.* (2004) The role of prophage-like elements in the diversity of *Salmonella enterica* serovars. *J. Mol. Biol.*, **339**, 279–300.
55. Kropinski,A.M., Sulakvelidze,A., Konczy,P. and Poppe,C. (2007) *Salmonella* phages and prophages–genomics and practical aspects. *Methods Mol. Biol.*, **394**, 133–175.
56. Lemire,S., Figueroa-Bossi,N. and Bossi,L. (2011) Bacteriophage crosstalk: coordination of prophage induction by trans-acting antirepressors. *PLoS Genet.*, **7**, e1002149.
57. Briani,F., Deho,G., Forti,F. and Ghisotti,D. (2001) The plasmid status of satellite bacteriophage P4. *Plasmid*, **45**, 1–17.
58. Padalon-Brauch,G., Hershberg,R., Elgrably-Weiss,M., Baruch,K., Rosenshine,I., Margalit,H. and 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.
59. Forti,F., Dragoni,I., Briani,F., Deho,G. and Ghisotti,D. (2002) Characterization of the small antisense CI RNA that regulates bacteriophage P4 immunity. *J. Mol. Biol.*, **315**, 541–549.
60. Sabbattini,P., Forti,F., Ghisotti,D. and Deho,G. (1995) Control of transcription termination by an RNA factor in bacteriophage P4 immunity: identification of the target sites. *J. Bacteriol.*, **177**, 1425–1434.
61. Navarre,W.W., Porwollik,S., Wang,Y., McClelland,M., Rosen,H., Libby,S.J. and Fang,F.C. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science*, **313**, 236–238.
62. Iyer,L.M., Babu,M.M. and Aravind,L. (2006) The HIRAN domain and recruitment of chromatin remodeling and repair activities to damaged DNA. *Cell Cycle*, **5**, 775–782.
63. Kuhle,V., Jackel,D. and Hensel,M. (2004) Effector proteins encoded by *Salmonella* pathogenicity island 2 interfere with the microtubule cytoskeleton after translocation into host cells. *Traffic*, **5**, 356–370.

64. Miao,E.A. and Miller,S.I. (2000) A conserved amino acid sequence directing intracellular type III secretion by *Salmonella* Typhimurium. *Proc. Natl Acad. Sci. USA*, **97**, 7539–7544.
65. Freeman,J.A., Ohl,M.E. and Miller,S.I. (2003) The *Salmonella enterica* serovar Typhimurium translocated effectors SseJ and SifB are targeted to the *Salmonella*-containing vacuole. *Infect. Immun.*, **71**, 418–427.
66. Lucchini,S., Rowley,G., Goldberg,M.D., Hurd,D., Harrison,M. and Hinton,J.C. (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.*, **2**, e81.
67. Manna,D., Porwollik,S., McClelland,M., Tan,R. and Higgins,N.P. (2007) Microarray analysis of Mu transposition in *Salmonella enterica*, serovar Typhimurium: transposon exclusion by high-density DNA binding proteins. *Mol. Microbiol.*, **66**, 315–328.
68. Hensel,M., Shea,J.E., Raupach,B., Monack,D., Falkow,S., Gleeson,C., Kubo,T. and Holden,D.W. (1997) Functional analysis of ssaJ and the ssaK/U operon, 13 genes encoding components of the type III secretion apparatus of *Salmonella* Pathogenicity Island 2. *Mol. Microbiol.*, **24**, 155–167.
69. Hensel,M., Shea,J.E., Waterman,S.R., Mundy,R., Nikolaus,T., Banks,G., Vazquez-Torres,A., Gleeson,C., Fang,F.C. and Holden,D.W. (1998) Genes encoding putative effector proteins of the type III secretion system of *Salmonella* pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. *Mol. Microbiol.*, **30**, 163–174.
70. Ohlson,M.B., Fluhr,K., Birmingham,C.L., Brumell,J.H. and Miller,S.I. (2005) SseJ deacylase activity by *Salmonella enterica* serovar Typhimurium promotes virulence in mice. *Infect. Immun.*, **73**, 6249–6259.
71. Doyle,M., Fookes,M., Ivens,A., Mangan,M.W., Wain,J. and Dorman,C.J. (2007) An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science*, **315**, 251–252.
72. Dillon,S.C., Cameron,A.D., Hokamp,K., Lucchini,S., Hinton,J.C. and Dorman,C.J. (2010) Genome-wide analysis of the H-NS and Sfh regulatory networks in *Salmonella* Typhimurium identifies a plasmid-encoded transcription silencing mechanism. *Mol. Microbiol.*, **76**, 1250–1265.
73. Rabsch,W., Voigt,W., Reissbrodt,R., Tsolis,R.M. and Baumler,A.J. (1999) *Salmonella* Typhimurium IroN and FepA proteins mediate uptake of enterobactin but differ in their specificity for other siderophores. *J. Bacteriol.*, **181**, 3610–3612.
74. Edwards,M.F. and Stocker,B.A. (1988) Construction of delta aroA his delta pur strains of *Salmonella* Typhi. *J. Bacteriol.*, **170**, 3991–3995.
75. Tsolis,R.M., Baumler,A.J., Heffron,F. and Stojiljkovic,I. (1996) Contribution of TonB- and Feo-mediated iron uptake to growth of *Salmonella* Typhimurium in the mouse. *Infect. Immun.*, **64**, 4549–4556.
76. Crick,F.H. (1966) Codon–anticodon pairing: the wobble hypothesis. *J. Mol. Biol.*, **19**, 548–555.
77. Lowe,T.M. and Eddy,S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic. Acids. Res.*, **25**, 955–964.
78. Nasvall,S.J., Chen,P. and Bjork,G.R. (2004) The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons in vivo. *RNA*, **10**, 1662–1673.
79. Osawa,S., Jukes,T.H., Watanabe,K. and Muto,A. (1992) Recent evidence for evolution of the genetic code. *Microbiol. Rev.*, **56**, 229–264.
80. Marck,C. and Grosjean,H. (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA*, **8**, 1189–1232.
81. Rosenblad,M.A., Larsen,N., Samuelsson,T. and Zwieb,C. (2009) Kinship in the SRP RNA family. *RNA Biol.*, **6**, 508–516.
82. Frank,D.N. and Pace,N.R. (1998) Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.*, **67**, 153–180.
83. Winkler,W., Nahvi,A. and Breaker,R.R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, **419**, 952–956.
84. Vogel,J., Bartels,V., Tang,T.H., Churakov,G., Slagter-Jager,J.G., Huttenhofer,A. and Wagner,E.G. (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.*, **31**, 6435–6443.
85. Tjaden,B., Saxena,R.M., Stolyar,S., Haynor,D.R., Kolker,E. and Rosenow,C. (2002) Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Res.*, **30**, 3732–3738.
86. Aseev,L.V., Levandovskaya,A.A., Tchufistova,L.S., Scaptsova,N.V. and Boni,I.V. (2008) A new regulatory circuit in ribosomal protein operons: S2-mediated control of the rpsB-tsf expression *in vivo*. *RNA*, **14**, 1882–1894.
87. Meyer,M.M., Ames,T.D., Smith,D.P., Weinberg,Z., Schwalbach,M.S., Giovannoni,S.J. and Breaker,R.R. (2009) Identification of candidate structured RNAs in the marine organism 'Candidatus *Pelagibacter ubique*'. *BMC Genomics*, **10**, 268.
88. Benard,L., Philippe,C., Ehresmann,B., Ehresmann,C. and Portier,C. (1996) Pseudoknot and translational control in the expression of the S15 ribosomal protein. *Biochimie*, **78**, 568–576.
89. Lease,R.A., Cusick,M.E. and Belfort,M. (1998) Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc. Natl Acad. Sci. USA*, **95**, 12456–12461.
90. Chao,Y., Papenfort,K., Reinhardt,R., Sharma,C.M. and 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.
91. Chen,S., Lesnik,E.A., Hall,T.A., Sampath,R., Griffey,R.H., Ecker,D.J. and Blyn,L.B. (2002) A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Biosystems*, **65**, 157–177.
92. Diwa,A., Bricker,A.L., Jain,C. and Belasco,J.G. (2000) An evolutionarily conserved RNA stem-loop functions as a sensor that directs feedback regulation of RNase E gene expression. *Genes Dev.*, **14**, 1249–1260.
93. Antal,M., Bordeau,V., Douchin,V. and Felden,B. (2005) A small bacterial RNA regulates a putative ABC transporter. *J. Biol. Chem.*, **280**, 7901–7908.
94. Wassarman,K.M., Repoila,F., Rosenow,C., Storz,G. and Gottesman,S. (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.*, **15**, 1637–1651.
95. Douchin,V., Bohn,C. and Bouloc,P. (2006) Down-regulation of porins by a small RNA bypasses the essentiality of the regulated intramembrane proteolysis protease RseP in *Escherichia coli*. *J. Biol. Chem.*, **281**, 12253–12259.
96. Rivas,E. and Eddy,S.R. (2001) Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics*, **2**, 8.
97. Kawano,M., Oshima,T., Kasai,H. and Mori,H. (2002) Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*. *Mol. Microbiol.*, **45**, 333–349.
98. Zurawski,G., Brown,K., Killingly,D. and Yanofsky,C. (1978) Nucleotide sequence of the leader region of the phenylalanine operon of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **75**, 4271–4275.
99. Naville,M. and Gautheret,D. (2010) Premature terminator analysis sheds light on a hidden world of bacterial transcriptional attenuation. *Genome Biol.*, **11**, R97.
100. Urban,J.H. and Vogel,J. (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate glmS mRNA translation. *PLoS Biol.*, **6**, e64.
101. Papenfort,K., Pfeiffer,V., Lucchini,S., Sonawane,A., Hinton,J.C. and Vogel,J. (2008) Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol. Microbiol.*, **68**, 890–906.
102. De Las Penas,A., Connolly,L. and Gross,C.A. (1997) SigmaE is an essential sigma factor in *Escherichia coli*. *J. Bacteriol.*, **179**, 6862–6864.
103. Papenfort,K., Pfeiffer,V., Mika,F., Lucchini,S., Hinton,J.C. and Vogel,J. (2006) SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Mol. Microbiol.*, **62**, 1674–1688.

104. Santiviago,C.A., Toro,C.S., Hidalgo,A.A., Youderian,P. and Mora,G.C. (2003) Global regulation of the *Salmonella enterica* serovar Typhimurium major porin, OmpD. *J. Bacteriol.*, **185**, 5901–5905.
105. Bossi,L. and Figueroa-Bossi,N. (2007) A small RNA downregulates LamB maltoporin in *Salmonella*. *Mol. Microbiol.*, **65**, 799–810.
106. Jones,A.M., Goodwill,A. and Elliott,T. (2006) Limited role for the DsrA and RprA regulatory RNAs in rpoS regulation in *Salmonella enterica*. *J. Bacteriol.*, **188**, 5077–5088.
107. Coornaert,A., Lu,A., Mandin,P., Springer,M., Gottesman,S. and Guillier,M. (2010) MicA sRNA links the PhoP regulon to cell envelope stress. *Mol. Microbiol.*, **76**, 467–479.
108. Griffin,J.E., Gawronski,J.D., Dejesus,M.A., Ioerger,T.R., Akerley,B.J. and Sassetti,C.M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.*, **7**, e1002251.
109. Gallagher,L.A., Shendure,J. and Manoil,C. (2011) Genome-scale identification of resistance functions in *Pseudomonas aeruginosa* using Tn-seq. *MBio*, **2**, e00315–e00310.
110. Gawronski,J.D., Wong,S.M., Giannoukos,G., Ward,D.V. and Akerley,B.J. (2009) Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc. Natl Acad. Sci. USA*, **106**, 16422–16427.
111. Citron,M. and Schuster,H. (1990) The c4 repressors of bacteriophages P1 and P7 are antisense RNAs. *Cell*, **62**, 591–598.
112. Ravin,N.V., Svarchevsky,A.N. and Deho,G. (1999) The anti-immunity system of phage-plasmid N15: identification of the antirepressor gene and its control by a small processed RNA. *Mol. Microbiol.*, **34**, 980–994.
113. Hobbs,E.C., Astarita,J.L. and Storz,G. (2010) Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: analysis of a bar-coded mutant collection. *J. Bacteriol.*, **192**, 59–67.