A transposon-derived small RNA regulates gene expression in *Salmonella* Typhimurium

Michael J. Ellis, Ryan S. Trussler, Onella Charles and David B. Haniford*

Department of Biochemistry, University of Western Ontario, London, Ontario N6A 5C1, Canada

Received September 21, 2016; Revised January 30, 2017; Editorial Decision January 31, 2017; Accepted February 06, 2017

ABSTRACT

Bacterial sRNAs play an important role in regulating many cellular processes including metabolism, outer membrane homeostasis and virulence. Although sRNAs were initially found in intergenic regions, there is emerging evidence that protein coding regions of the genome are a rich reservoir of sRNAs. Here we report that the 5'UTR of IS200 transposase mRNA (tnpA) is processed to produce regulatory RNAs that affect expression of over 70 genes in *Salmonella* Typhimurium. We provide evidence that the tnpA derived sRNA base-pairs with invF mRNA to repress expression. As InvF is a transcriptional activator of SPI-1 encoded and other effector proteins, tnpA indirectly represses these genes. We show that deletion of IS200 elements in *S.* Typhimurium increases invasion *in vitro* and reduces growth rate, while over-expression of tnpA suppresses invasion. Our work indicates that tnpA acts as an sRNA ‘sponge’ that sets a threshold for activation of *Salmonella* pathogenicity island (SPI)-1 effector proteins and identifies a new class of ‘passenger gene’ for bacterial transposons, providing the first example of a bacterial transposon producing a regulatory RNA that controls host gene expression.

INTRODUCTION

IS200 is the smallest prokaryotic transposon and is widely conserved in Enterobacteriaceae and found throughout *Enterobacter* and *Archaeeae*. One unusual feature of IS200 elements is the high copy number achieved in *Salmonella* spp. (1–3). Many strains of *Yersinia pestis* contain more than 50 copies of the IS200 ortholog IS1541, while strains of *Salmonella enterica* subsp. *enterica* serovar Typhimurium (*S.* Typhimurium) typically contain 5–12 copies and *S.* Typhi contains 26 copies of IS200 per genome. In the above cases, all IS200 paralogs appear to be 100% conserved and in general IS200 orthologs share >90% identity. A highly active transposon might be expected to achieve this high copy number and repeated transposition would maintain sequence identity of paralogs; however, IS200 is an essentially dormant transposon (1,2). Conservation and copy number might therefore reflect a selective pressure on the host bacterium to maintain IS200. Transposons can contribute to host fitness in several ways including: (i) by mediating DNA rearrangements that influence host gene expression and gene structure (4); (ii) contributing passenger genes such as antibiotic resistance determinants (5); (iii) providing a rich source of DNA regulatory sequences (6); (iv) providing proteins and/or protein motifs from transposase proteins that can be domesticated by the host (7); and (v) providing regulatory RNAs that affect host gene expression (8,9). As a simple insertion sequence, IS200 does not encode any passenger genes, and the dormancy of IS200 suggests that this element would not contribute transposition-dependent functions to the host.

Non-coding RNAs (ncRNA) play a crucial role in regulating many critical processes in bacteria including outer membrane homeostasis, metabolism and virulence (10,11). The largest class of bacterial ncRNA are small RNAs (sRNA) that base-pair with target mRNAs and affect translation and/or transcript stability. sRNAs are typically expressed from intergenic regions and therefore have limited sequence complementarity with their trans-encoded targets. A related class of ncRNA are antisense RNAs (asRNA) which are encoded on the opposite strand of DNA to their target mRNA. Accordingly, asRNAs have much more extensive complementarity with their cis-encoded targets. Note that both sRNAs and asRNAs act by an antisense mechanism, but are classified based on their genomic context relative to target mRNAs. The third and smallest class of bacterial ncRNAs act by binding to and regulating protein activity (e.g. 6S RNA, CsrB/C). The classic distinction between these three classes of ncRNA has been challenged with continually emerging examples of dual-function ncRNA, including sRNAs derived from mRNAs (12–15), base-pairing sRNAs acting to modulate protein activity (16–19) and asRNAs acting in trans to regulate genes expressed from different loci (20,21). One common feature for base-pairing ncRNAs is that the RNA-binding protein Hfq is typically required to facilitate pairing when there is limited complementarity between an sRNA and mRNA (22). In general, an interaction between Hfq and an...
This regulatory element ensures that impinging transcription does not activate\textit{tnpA} expression and terminates $\sim$85\% of upstream transcripts (23). Second, translation of\textit{tnpA} is strongly repressed by mRNA secondary structure that includes the Shine-Dalgarno sequence (SD) (Figure 1B). This stem-loop element represses\textit{tnpA} expression 20-fold by preventing 30S ribosome binding. Third, art200 base-pairs with\textit{tnpA} to inhibit ribosome binding, and reduces translation 15-fold. Lastly,\textit{tnpA} translation is inhibited directly by the RNA-binding protein Hfq, which recognizes a sequence immediately upstream of the SD and accordingly sterically occludes ribosome binding. The three post-transcriptional mechanisms act independently and together suppress translation of\textit{tnpA} by at least 750-fold, ensuring almost no TnpA protein is produced (1). While these regulatory mechanisms appear to be redundant,\textit{tnpA} expression is reasonably high in\textit{S. Typhimurium} for a transposon ($\sim$10\% the expression of\textit{hfq} in mid-exponential phase (25)). It therefore appears that IS200 elements have evolved to maintain moderate transcription of\textit{tnpA} from an IS200 encoded promoter, but close to no synthesis of TnpA. Another noteworthy feature of IS200-encoded RNAs is that art200 expression appears to be growth phase regulated, with increased expression when\textit{S. Typhimurium} transitions to stationary phase in rich media, as well as in growth media that stimulate\textit{Salmonella} pathogenicity island (SPI) expression (Supplementary Figure S1; (24)). Additionally, art200 interacts with Hfq \textit{in vivo}, although Hfq is dispensable for antisense regulation of\textit{tnpA} expression. Intriguingly, while art200 expression is increased in stationary-phase,\textit{tnpA} expression decreases $\sim$5-fold (25), which may indicate that art200 expression is altered to control\textit{tnpA} RNA levels. One explanation for the unusual characteristics of IS200-encoded RNAs is that a moderately expressed but never translated\textit{tnpA} provides a way in which IS200 transposition could be rapidly activated under certain conditions. However, previous work found that IS200 transposition is remarkably rare, even when post-transcriptional regulation is completely eliminated (1). With respect to art200’s expression patterns and Hfq-binding properties, this could simply reflect stochastic evolution of the promoter and sequence of a regulatory RNA. A more intriguing explanation for the peculiar properties of\textit{tnpA} and art200 is that one or both IS200-encoded RNAs serves a regulatory role independent of controlling transposition. In this scenario, an IS200 encoded RNA might provide a selective advantage to\textit{Salmonella} spp. and accordingly explain the conservation and high copy number of this transposon.

In the current work we performed an RNA-Seq experiment to ask if IS200-encoded RNAs affect gene expression in\textit{S. Typhimurium}. We provide evidence that the 5’ UTR of\textit{tnpA} represses many genes including the SPI-1 encoded transcription factor,\textit{invF}. Our data suggests that\textit{tnpA} base-pairs with\textit{invF}, and the consequence of this interaction is downregulation of the SPI-1 translocon (\textit{sicAsipBC}) and SPI-1 mediated invasion. This work is the first demonstration of a bacterial transposon encoding regulatory RNAs that influence host gene expression.

\section*{Figure 1.} IS200 and experimental approach. (A) IS200 encodes a transposase mRNA (\textit{tnpA}, red) and an antisense RNA (art200, blue). The\textit{tnpA\textsubscript{trunc}-255} transcript encodes the first 255 nt of\textit{tnpA} fused to the last 108 nt of SgrS (black, includes an intrinsic terminator) and is expressed from the Tet promoter. (B) Approach used to deplete art200. Pairing between\textit{tnpA} (red) and art200 (blue) results in degradation of art200. The M1 mutation alters three critical nucleotides in the terminal loop of\textit{tnpA} and prevents pairing with art200. The sequenced Shine-Dalgarno sequence (SD) of\textit{tnpA} is indicated with a box and the translation start codon (AUG) is shown. (C) Heat map showing expected expression of IS200 RNAs in\textit{Salmonella Typhimurium} LT2 containing plasmids over-expressing either wild-type (WT) or M1 forms of\textit{tnpA\textsubscript{trunc}-255}. Note that ‘\textit{tnpA}’ signifies endogenous transposase transcript and ‘vector’ is a control plasmid that does not encode\textit{tnpA}.

mRNA or sRNA indicates that the RNA is involved in post-transcriptional regulation via a base-pairing mechanism.

IS200 elements express two RNA molecules (Figure 1A), the first is an mRNA encoding the transposase protein (\textit{tnpA}), and the second is an asRNA (art200, previously named STnc490) that is complementary to the\textit{tnpA} 5’UTR (1,23,24). Expression of the IS200 TnpA is strongly repressed by four independent mechanisms. First, the left-end of IS200 contains an inverted repeat that forms a strong, bidirectional, Rho-independent transcriptional terminator.
MATERIALS AND METHODS

Growth conditions, strains and plasmids

Unless otherwise stated, S. Typhimurium was grown at 37°C with shaking in Lennox Broth (LB; 5 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract). For experiments where RNA was extracted at multiple time-points, overnight cultures were diluted once (1:100 into 7 or 25 ml) and aliquoted (2 ml) into separate culture tubes for each time point. For SPI-1 inducing conditions, cells were grown as previously described (24). For SPI-2 inducing conditions, cells were grown overnight in LB and diluted 1:100 into acidic low-phosphate, low-magnesium media (80 mM MES pH 5.8, 5 mM KCl, 7.5 mM (NH₄)₂SO₄, 0.5 mM K₂SO₄, 38 mM glycerol, 0.1% casamino acids [w/v], 8 μM MgCl₂, 337 μM KH₂PO₄) (26). Where appropriate, antibiotics were used at the following concentrations: tetracycline (tet), 15 μg/ml; chloramphenicol, 20 μg/ml; kanamycin (kan), 25 μg/ml; streptomycin (str), 150 μg/ml. For experiments with marked alleles, selection was only used in the overnight culture.

All strains and plasmids used in this study are listed in Table S4 and oligonucleotides are listed in Table S5. S. Typhimurium str. LT2 or SL1344 were considered wild-type (WT) strains, and derivative strains were made in the ΔtnpA (DBH415) backbone. Escherichia coli DH5α was used for routine cloning and plasmid propagation.

Mutant strains of SL1344 were constructed by Lambda Red recombineering (27) and all mutations were checked by colony Polymerase chain reaction (PCR). DBH401 (ΔtnpA4/6/7) and DBH415 (ΔtnpA1-7, referred to as ΔtnpA) were constructed by transducing individual IS200 knockout alleles into a single strain. DBH393 and related strains were created by inserting a kan-pTet (or cm-pTet for DBH398) cassette in front of tnpA such that the Tet promoter is driving transcription of tnpA (tnpA7::kan-pTet). Complementation strains were constructed by transducing the tnpA7::kan-pTet or tnpA7::kan-pTet(+19) from DBH416 or DBH419 into the ΔtnpA (DBH415) background. Further details of strain and plasmid construction are provided in Supplementary Materials and Methods.

RNA isolation, northern blot and primer extension

Total RNA was prepared by the hot acid phenol method (28). Northern blots were performed as previously described (19) using 5 or 10 μg of total RNA and 532P-labeled oligonucleotide probes (oDH428 tnpA; oDH427, art200) or a uniformly 32P-labeled riboprobe (5S rRNA, generated with oDH234 and oDH235; art200, generated with oDH450 and oDH394). Primer extension was performed as previously described (1) using 9 μg of total RNA and primers oDH428 or oDH394 (tnpA) or oDH710 (invF). Processed RNA was eliminated by terminator exonuclease (TEX) (Epitome) treatment according to the manufacturer’s instructions.

RNA-seq and data analysis

Salmonella Typhimurium LT2 was transformed with pDH900 (empty vector), pDH899 (pTet-tnpAWT-255) or pDH914 (pTet-tnpAtruncML-255). Two colonies from each transformation were each used to inoculate 1 ml of LB-Luria (0.5 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract) with tet and were grown for 8 h. Precultures were subcultured 1:100 into LB-Luria and grown for 16 h. Total RNA was isolated and treated with TURBO DNase (Ambion) to remove residual genomic DNA and submitted to the London Regional Genomic Centre for library preparation and sequencing. Libraries were prepared with the Riboprobe Zero (Gram-Negative Bacteria) (Epitome) and ScriptSeq v2 (Epitome) kits. The six libraries (two biological replicates from each strain) were pooled and sequenced with 50 cycles on an Illumina MiSeq. Reads were aligned to the S. Typhimurium LT2 genome (NC_003197) with Rockhopper (29) (Table S1) and differential expression was analyzed using ALDEX2 (30). More detail on data analysis is provided in Supplementary Materials and Methods.

Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

DNase treated RNA (2 μg) was converted to cDNA with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems); cDNA was diluted to 30 ng/μl in TE (50 mM Tris–HCl, pH 8.0, 1 mM Ethylenediaminetetraacetic acid (EDTA)) and stored at −20°C. A minimum of three biological replicates were analyzed in technical triplicate in each experiment and the 16S rRNA (rrsA) was used as a reference gene for relative quantitation. Reactions (20 μl) contained 10 ng of cDNA, 500 nM of each primer (Supplementary Table S5) and PowerUP SYBR Green Master Mix (Applied Biosystems). Standard settings on the Viia 7 Real-Time PCR System were used except for the anneal/extension step, which was performed at 60.5°C. Relative expression of each target was calculated by the efficiency corrected method (31). The amplification efficiency was determined for tnpA (2.20), thrS (2.04), rrsA (2.00), invF (2.12), sipB (2.03), sipC (2.01) and sipA (2.00); an efficiency of 2.0 was used for all other primer pairs.

Western blot

DBH388 (invF::3X-FLAG-kan) transformed with pDH900 (empty vector), pDH960 (pTet-tnpAtrunc-50) or pDH962 (pTet-tnpAtrunc-200) was grown to OD₆₀₀ = 0.5 and cells from 1 ml of culture was collected by centrifugation. For the experiment comparing DBH388 to DBH398 (invF::3X-FLAG-kan tnpA7::kan-pTet), the volume of culture was adjusted so that an equivalent of 0.5 OD were harvested. The cell pellet was resuspended in 200 μl of sodium dodecyl sulphate (SDS) sample buffer (60 mM Tris–HCl, pH 6.8, 2% SDS [w/v], 0.01% bromophenol blue [w/v], 1% β-mercaptoethanol [v/v]) and boiled for 5 min. Samples (10 μl) were resolved on 10% polyacrylamide gels and electroblotted to a polyvinylidene difluoride (PVDF) membrane. Membranes were incubated in 5% milk overnight with primary antibody (1:5000 dilution: mouse α-FLAG M2, Sigma; rabbit α-GroES, Sigma; mouse α-DnaK, Enzo), followed by incubation with a 1:5000 dilution of secondary antibody (α-mouse-HRP or α-rabbit-HRP, Promega). Blots were developed with a Pierce ECL 2.
western blotting substrate and a STORM scanner. Membranes were stripped and re-probed for loading controls (GroES/DnaK). Bands were quantitated in ImageQuant and the amount of InvF-3× FLAG was normalized to the internal standard (GroES/DnaK) and then the control strain (empty vector or DBH388).

Gentamicin protection (invasion) assay

Invasion assays were performed essentially as previously described (32). Tissue-culture plates (24-well) were seeded with ~0.05 × 10⁶ HeLa cells per well in 1 ml of Dulbecco’s modified eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS), penicillin (100 U/ml) and streptomycin (100 μg/ml) 20–22 h prior to the invasion assay. At the time of the assay cells were 60–70% confluent (~0.1 × 10⁶ cells per well).

Freshly streaked colonies of DBH347 (SL1344 WT), DBH393 (SL1344 ΔtnpA) or DBH415 (SL1344 ΔtnpA Δart200) were used to inoculate 2 ml of SPI-1 inducing media (17.5 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract) containing 150 μg/ml streptomycin and 25 μg/ml kan (for DBH393 and DBH418). Overnight cultures were subcultured 1:100 into LB and grown for 2 h or 3 h with shaking to mid-log (OD₆₀₀ = 0.5) or late-exponential phase (OD₆₀₀ = 1.2). Bacterial cells were washed with Phosphate buffered saline (PBS) and diluted in DMEM/10% FBS to a concentration of 1 × 10⁹ cfu/ml.

HeLa cells were washed with PBS and 1 ml of bacterial suspension (MOI of 100) was added to two wells for each culture (technical duplicate). Serial dilutions of the bacterial suspension were plated on LB agar plates with 150 μg/ml streptomycin to determine the input number of bacteria.

Bacterial cells were centrifuged onto the HeLa monolayer at 500 × g for 3 min at room temperature and then incubated at 37°C for 10 min. Bacterial cells were washed away with PBS and 1 ml of fresh DMEM/10% FBS was added to each well, followed by a 20 min incubation at 37°C. Culture media was replaced with DMEM/10% FBS containing 100 μg/ml gentamicin followed by a 30 min incubation at 37°C to kill extracellular bacteria. After washing with PBS, HeLa cells were resuspended in 1 ml lysis solution (PBS, 5 mM EDTA, 0.5% Triton X-100, 0.1% SDS) and serial dilutions were plated on LB with 150 μg/ml streptomycin to determine the output bacterial cell counts. Invasion was calculated as the ratio of recovered cells to the input and normalized to the WT strain for each experiment.

Electrophoretic mobility shift assay (EMSA) and lead footprinting

In vitro pairing experiments were performed as previously described (1,33) except that the RNAs were mixed prior to denaturation.

Growth curves

Growth was measured in a Multiskan Go microplate spectrophotometer. Cells from two overnight cultures (biological replicates) of each strain (DBH347, WT; DBH415, ΔtnpA; DBH416, ΔtnpA/tnpA Δart200 Δkan-pTet) were washed with sterile saline and diluted 100-fold into LB. Two hundred microliters of each dilution was added to three wells (technical replicates) of a 96-well microplate. Cultures were grown with continuous shaking at 37°C for 12 h and absorbance at 600 nm (A₆₀₀) was measured every 15 min. Note that the A₆₀₀ was not adjusted for path length and light scattering from the microplate lid and is therefore not directly comparable to optical density readings measured in a standard cuvette.

RESULTS

Profiling changes in S. Typhimurium gene expression in response to altered levels of IS200-encoded transcripts

We used RNA-Seq to analyze gene expression in S. Typhimurium LT2 under conditions where levels of tnpA and art200 were altered from native levels. In one strain we introduced a plasmid that constitutively over-expresses a truncated form (nt 1–255) of the transposase mRNA (tnpAΔ truncWT, Figure 1A). This strain produces very low amounts of art200 because tnpAΔ truncWT-255 RNA pairs with art200 and this pairing promotes degradation of art200 (Figure 1B and C; Supplementary Figure S1C). When we looked for differentially expressed genes in this strain versus an empty vector control strain, we identified 187 genes with altered expression (Figure 2A, black dots; Supplementary Table S3), 99 of which had at least a 2-fold change in expression. This altered pattern of gene expression could arise from either depletion of art200 and/or the over-expression of the truncated tnpA mRNA. To distinguish between these possibilities, we profiled gene expression in a third strain expressing a truncated form of tnpA (tnpAΔ truncM1-255) that is unable to pair with art200 (Figure 2B). Genes affected by depletion of art200 would show differential expression when tnpAΔ truncM1-255 was over-expressed but not when tnpAΔ truncWT-255 was over-expressed. When all three comparisons were made, only six genes appeared to be uniquely regulated by art200 (Figure 2C; glnH, gltI, acs, icdA, hutU and a predicted asRNA to the 3′ end of fadR). In contrast, genes regulated by tnpAΔ trunc over-expression would show differential expression in both WT and M1 tnpAΔ trunc strains when compared to an empty vector. A total of 73 genes fit this criterion (Figure 2C). Based on this analysis we concluded that transcripts derived from IS200 impact on host gene expression and that high levels of a truncated form of tnpA that includes the 5′ UTR has a greater impact on host gene expression than depletion of art200.

Lastly, we searched for cellular processes enriched with genes affected by tnpAΔ truncWT over-expression. This analysis found that tnpA over-expression significantly represses genes involved in pathogenesis, glycerol-3-phosphate metabolism and oxidation–reduction reactions (Figure 2D). Similar results were obtained when pathway analysis was performed on the 73 genes affected by both WT and M1 tnpAΔ trunc constructs (Supplementary Figure S1D). The strongest change in gene expression in any of these pathways was the SPI-1 encoded effector protein, SpotC (10-fold repression by over-expression of tnpAΔ truncWT-255). As S. Typhimurium LT2 is avirulent (34), we switched to the virulent SL1344 strain (seven copies of IS200 versus six copies of IS200 in LT2) for subsequent studies.
Figure 2. Summary of RNA-Seq data. (A and B) Expression plot comparing relative abundance (log2 clr) of *Salmonella Typhimurium* LT2 transcripts in the presence of an empty vector (x-axis) or plasmid expressing WT (y-axis, A) or M1 (y-axis, B) *tnpA*trunc-255. Differentially expressed genes (Effect size >2) are indicated in black and dotted lines indicate a 2-fold change in expression from the line of best fit for the data (A, Pearson’s r = 0.9393; B, Pearson’s r = 0.9348). Reads derived from *tnpA*trunc-255 mapped to either the IS200 transposase coding sequence (*tnpA*) or 5′ UTR (*tnpA*5′ UTR) and are indicated in blue. SPI-1 genes *sicA*, *sipB*, *sipC* and *invF* are highlighted in red; note that *invF* was repressed >3-fold by *tnpA*truncWT-255 but fell below our cut-off for differential expression (Effect size = −1.2025). Genes with an Effect size <2 are indicated in grey and are not considered to be differentially expressed. (C) Venn diagram showing the overlap of genes identified as differentially expressed when comparing the empty vector to *tnpA*truncWT-255 (blue) or *tnpA*truncM1-255 (yellow) or *tnpA*truncWT-255 to *tnpA*truncM1-255 (green). (D) Results of GO Enrichment Analysis. The 187 genes identified as differentially expressed when comparing expression in the presence of the vector versus *tnpA*truncWT-255 were used as a query gene list for GO Enrichment Analysis. The log2 fold change (Vector versus *tnpA*truncWT-255) of genes in the three enriched biological processes are shown along with the enrichment score and P-value from the PANTHER Over-representation test. Horizontal bars indicate the median fold-change for each biological process.

Characterization of *tnpA* derived RNAs

Our RNA-Seq analysis revealed that over-expression of the first third of transposase mRNA had a substantial impact on gene expression in *S. Typhimurium*. While this points to *tnpA* mRNA acting as a regulatory RNA, we thought it more likely that a naturally truncated or processed form of *tnpA* is produced from the 5′ end to act as a regulatory RNA. This would be in line with other recently discovered mRNA derived sRNAs (35). We initially looked for evidence of an sRNA derived from the 5′ end by performing a northern blot (5′ UTR probe) on RNA isolated from a strain expressing native levels of *tnpA* (WT) or a strain where *tnpA* was over-expressed through the fusion of the pTet promoter to one copy of *tnpA* in the chromosome. In the latter strain we detected three species, two of which are ~90 and ~110 nt and the other is >310 nt (Figure 3A, lane 3). The 90 and >310 nt species were also just detectable in the strain expressing *tnpA* at native levels (lane 1). In contrast, none of these species were detected in a strain where four of seven copies of the *tnpA* gene were deleted (lane 2). Additionally, both the 110- and 90-nt species were detected by northern blots on samples where *tnpA*trunc-255 was over-expressed (Figure 4A). Taken together these results show that: (i) the native *tnpA* gene generated one or more sRNAs; (ii) sRNA production does not require more than 255 bp of the *tnpA* gene; and (iii) sRNA production occurs independent of the promoter used to drive *tnpA* transcription. The latter point is suggestive of sRNAs being produced through RNA processing of the *tnpA* transcript.

We next performed primer extension on the above RNA samples to map 5′ ends of each species. In one experiment we used a primer that anneals to the 5′ UTR (nt 46–64). The results show that the majority of *tnpA* transcripts start at
Figure 3. Processing of the tnpA transcript. (A) A northern blot of tnpA RNA isolated from SL1344 strains expressing tnpA at endogenous levels (WT), over-expressing tnpA from the tnpA,7 locus (tnpA,7::kan-pTet, pTet) or with a reduced number of endogenous copies of tnpA (ΔtnpA2/4/6/7, ΔtnpA). Full-length (closed circle) and processed (open circle) forms of tnpA were detected with a probe that anneals to the tnpA 5'UTR (oDH429). 5S rRNA was used as a loading control. (B and D) 5' ends of tnpA were mapped using primer extension. RNA was isolated from the above strains (two replicates) and tnpA was detected using a primer that anneals to the 5'UTR (nt 46-64, B) or coding sequence (nt 151-171, D). ddNTP sequencing lanes (using tnpA,7::kan-pTet RNA as a template) were used to determine the nucleotide position of primer extension products relative to the transcription start site (+1, 'Full-length'). (C) tnpA is processed at U17 and A19. RNA isolated from the WT or tnpA,7::kan-pTet strains was treated with TEX (+) or incubated with buffer (-) before tnpA was detected by primer extension. (E) Summary of primer extension experiments. The major primer extension products from parts B and D are illustrated along with the primer binding sites. The two primers used for primer extension would detect different molecules of tnpA based on processing occurring between the primer binding sites. From the positions of 5' ends and the size of low molecular weight RNA species in the northern (part A), we infer that the tnpA transcript is processed at two sites to produce two stable 5' UTR-containing species (site B, tnpA-110; sites A+B, tnpA-90). (F) Proposed processing pathway for tnpA. Full-length tnpA would be processed at site 'B' (purple) generating tnpA-110. Subsequent processing at site 'A' on tnpA (red) generates tnpA-90, which is the most stable tnpA species. The binding site for the northern probe (oDH429) is indicated in green.
position 19 rather than the expected transcription start site (Figure 3B). This pattern was observed both when tnpA was over-expressed and activated at native levels. We also show that prior treatment of the RNA with 5′monophosphate-dependent TEX resulted in loss of the primer extension signal at nt 19, indicating that this 5′end is generated through transcript processing (Figure 3C). In a second experiment we used a primer that anneals in the coding sequence (nt 151–171) (Figure 3D). Here we also identified the position 19 5′end and additional 5′ends surrounding position 108. These alternative 5′ends were also lost upon TEX treatment, indicating processing in a second region of the tnpA transcript (Supplementary Figure S2A). Processing events at positions 19 and 108 would generate a 5′UTR containing species of ~90 nt. In contrast, processing at only the downstream site would generate a 5′UTR containing species of ~110 nt in length. Based on these experiments, we infer that processing at sites designated A and B in Figure 3D generates stable tnpA encoded sRNAs (Figure 3E and F).

Repression of SPI-1 encoded genes by tnpA

To test the hypothesis that one or both of the above described sRNAs are actually the active molecules for regulating host genes, we made additional tnpA_{trunc} constructs (first 50, 200 and 250 nt of tnpA over-expressed from plasmids, Supplementary Figure S3A) to determine the mini-
nal tnpA required for affecting gene expression in S. Typhimurium; both tnpA<sub>trunc</sub>=−200 and −250 are processed to produce ~110 and ~90 nt species (Figure 4A). Note that tnpA<sub>trunc</sub>=−50 is not detected by northern blot as this construct does not contain the sequence recognized by the northern probe. We used RT-qPCR to determine which of these truncated tnpA molecules downregulates a set of functionally related genes (sicA, sipB and sipC) identified in our RNA-Seq experiment to be repressed by tnpA. All three truncated forms of tnpA downregulated sicA, sipB and sipC expression (>2.5-fold) but not the expression of thrS, a gene whose expression was not affected by tnpA in the RNA-Seq analysis (Figure 4B). From this experiment it is evident that over-expression of only the first 50 nt of tnpA is sufficient to negatively regulate expression of the aforementioned genes, indicating that either the 110- or 90-nt processed species is a functional sRNA. It may also be significant that of the three truncated forms of tnpA tested in this experiment, tnpA<sub>trunc</sub>=50 downregulated expression of the target genes to the highest degree and is the only one of the three tnpA RNAs incapable of base-pairing with art200 (Figure 4A). Indeed, all three truncated forms of tnpA on sicAsipBC could therefore be mediated through direct regulation of invF. Indeed, all three tnpA<sub>trunc</sub> constructs repressed invF, with over-expression of tnpA<sub>trunc</sub>=−50 reducing invF mRNA levels 3.5-fold (Figure 4B). We also examined the effect of constitutive over-expression of tnpA<sub>trunc</sub> on InvF protein levels with a strain of SL1344 containing a 3× FLAG tag integrated at the C-terminus of the native invF gene. Consistent with our RT-qPCR analysis, tnpA<sub>trunc</sub> repressed InvF protein levels over 2-fold (Figure 4C).

We also looked at the ability of tnpA to inhibit invF expression using over-expressed full-length tnpA (tnpA<sub>7</sub>::kan-pTet, Supplementary Figure S3B). We show that in late-exponential phase this strain expressed tnpA at a level ~65-fold higher than the WT strain, and decreased invF transcript and protein levels 2-2.5-fold (Figure 4D and E). For comparison, invF levels were decreased 5.7-fold in a ΔhilA strain. As HilA is a transcriptional activator of invF, the ΔhilA strain provides a measure of uninduced invF expression. Together, the above data indicates that a tnpA-derived sRNA inhibits expression of SPI-1 effectors proteins SicA, SipB, SipC by repressing InvF expression.

Salmonella Typhimurium employs the SPI-1 T3SS for crossing the intestinal epithelium during the course of an oral infection. Our data thus far shows that tnpA over-expression represses expression of components of the SPI-1 T3SS, and we therefore asked if tnpA affects invasion of non-phagocytic cells in vitro. We infected cultured HeLa cells with WT, tnpA<sub>7</sub>::kan-pTet or non-invasive (ΔinvA) strains of SL1344 to determine if over-expression of full-length tnpA alters the rate of invasion. To separate initial invasion from intracellular replication, we used a short time of infection (10 + 30 min recovery) before killing extracellular bacterial cells. As shown in Figure 4F, over-expression of tnpA reduces invasion 2-fold relative to the WT strain. The agreement between our expression data and invasion experiments led us to conclude that tnpA represses SPI-1 mediated invasion of non-phagocytic cells, likely through inhibition of invF expression.

Direct interaction between tnpA and invF

Our work thus far indicated that the tnpA mRNA is processed to produce a non-coding RNA (tnpA-90 and/or tnpA-110) that represses invF. As many bacterial ncRNAs act by base-pairing mechanisms, we first used IntaRNA (38) to find predicted base-pairing interactions between the 5′ end of tnpA and invF. We identified a single extended region of predicted complementarity between the first 63 nt of tnpA and an interval 104–160 nt upstream of the start codon on invF (Figure 5A). This predicted interaction fits with the above data showing that the first 50 nt of tnpA is sufficient for repressing invF, and supports tnpA-90 or -110 acting as an sRNA. We used a gel shift assay to determine if tnpA and the 5′ end of invF can base-pair in vitro. As the reported transcription start site (+1, TSS) for invF is 132-nt upstream of the start codon (in the center of the predicted pairing region) (39), we elected to start the in vitro transcript for invF at this position. We observed a modest shift in 32P-labeled invF upon incubation with increasing concentrations of unlabeled tnpA (first 173 nt) (Figure 5B, lanes 1–4). Importantly, a complex of the same mobility formed when 32P-labeled tnpA was incubated with unlabeled invF (lanes 5–8). To determine the specificity of tnpA:invF pairing, we assayed the ability of a previously characterized mutant form of tnpA (tnpA<sup>A</sup>S, Supplementary Figure S3C, (1)) to pair with invF; pairing was mostly lost as a consequence of the LS mutations in tnpA (lanes 9–11).

We next used Pb<sup>2+</sup> footprinting to define the region on invF that base-pairs with tnpA. 5′ 32P-labeled invF was incubated with a 5- or 10-fold excess of tnpA (WT or LS) before the addition of Pb(II)-acetate. The most substantial region of pairing was a 7-nt interval located 17–23 nt after the invF TSS (lanes 3–5, Figure 5A and C).

To test if this interaction occurs in vivo, we introduced mutations into the tnpA<sub>7</sub>::kan-pTet construct (T1 mutations, Supplementary Figure S3C) that prevent base-pairing with nt 17–23 of invF (Supplementary Figure S4). We performed RT-qPCR on RNA extracted from SL1344 WT, tnpA<sub>7</sub>::kan-pTet and tnpA<sub>7</sub>::kan-pTet-T1 strains grown to late-exponential phase. Over-expression of the WT tnpA reduced invF and sicA levels 3.5- and 2-fold respectively, while the T1 mutant form of tnpA did not affect either of these transcripts (Figure 5D). Due to the complex transcriptional regulation of invF and the location of the pairing region (~20 nt downstream of the TSS) we have not introduced compensatory mutations to invF. This experiment showed that the effect of tnpA on SPI-1 expression is sequence specific; combined with our <i>in vitro</i> pairing experiments, the above data is consistent with the 5′ ends of tnpA.
Figure 5. Evidence for a base-pairing interaction between invF and tnpA. (A) Predicted pairing interaction between the first 63 nt of tnpA and a region of invF 104–160 nt upstream of the start codon. Note that the main transcription start site (TSS, +1) for invF is 132 nt upstream of the start codon and nucleotides upstream of the TSS are shown in grey. invF nucleotides shown experimentally to be involved in pairing with tnpA are indicated in red; tnpA LS and T1 mutations are shown in bold. (B) Pairing between tnpA and invF was measured by electrophoretic mobility shift assay. $^{32}$P-labeled invF (−132 to +66 relative to the start codon) or tnpA (−103 to +71 relative to the start codon) was incubated with increasing concentrations of unlabeled tnpA or invF respectively (labeled RNA, 2.4 nM; unlabeled RNA 24, 120, 240 nM) and pairing reactions were analyzed by native PAGE. A mutant form of tnpA (tnpALS) was also included in this experiment. Certain lanes have been removed from one gel for clarity (vertical white line separating lanes 8 and 9). Reactions containing only the labeled RNA (lanes 1, 5 and 9) are indicated with ‘−’. (C) Pb$^{2+}$ footprinting was used to analyze base-pairing between $^5$'$^3$2P-labeled invF (70 nM) and unlabeled tnpAWT or tnpALS (same transcripts as in B). An RNase T1 sequencing reaction (G, lane 1) was used to assign positions of lead sensitivity (numbers relative to the 5' end), and an untreated RNA control (UT, lane 2) is shown. Red bars to the right of the gel image highlight tnpA$^{WT}$–dependent protections on invF. (D) RT-qPCR from RNA isolated from the indicated SL1344 strains grown to late-exponential phase (OD$_{600}$ = 1.4). Error bars show standard error on the mean (n = 4).
and invF base-pairing, the consequence of which is reduced invF mRNA levels. The ultimate test of this model would be to introduce mutations to invF to restore complementarity with the T1 mutant of tnpA.

Over-expression of tnpA represses expression of SPI-1 in a growth phase dependent manner

We next asked if the regulation of SPI-1 genes by a tnpA-derived sRNA is linked to growth phase, as invF expression is induced in late exponential and early stationary phase.

We profiled the expression of invF and other SPI-1 encoded genes (sicA, sipB, sipC and prgH) during five different growth phases in the WT or tnpA over-expression (tnpA::kan-pTet) strains. Importantly, there was no difference in growth rate between the two strains (Supplementary Figure S5A). Over-expression of tnpA did not affect SPI-1 gene expression in cells in lag- or early-exponential phase (Figure 6A and B). In both of these growth phases, tnpA in the WT strain was expressed higher than invF (Figure 6F), suggesting that the native expression of tnpA was sufficient for fully repressing invF. Once cells reached late-exponential phase, over-expression of tnpA repressed invF (2-fold), sicA (5.5-fold), sipB (4-fold) and sipC (2-fold); prgH expression (an InvF-independent SPI-1 encoded gene) was not affected by tnpA over-expression (Figure 6C). At this growth phase invF is moderately induced (~2-fold) relative to early-exponential phase, and is now present at ~2-fold excess to tnpA in the WT strain (Figure 6F). Here, endogenous tnpA would be limiting, explaining why this growth phase shows the largest impact of tnpA over-expression. Lastly, tnpA over-expression had a subtle effect on invF expression during early- and deep-stationary phase growth, which is likely due to the high expression of invF relative to tnpA (Figure 6D–F).

Together, these data show that tnpA over-expression affects invF levels only when native tnpA is expressed at lower levels than invF. This suggests the stoichiometry between both transcripts is important, and is consistent with a direct interaction between tnpA and invF. Additionally, the growth phases where tnpA over-expression repressed sicAsipBC were the same as those where tnpA repressed invF, providing additional support to a model where tnpA acts through invF to repress sicAsipBC.

Contribution of native tnpA expression to the regulation of SPI-1 expression

The observation that over-expressing tnpA only affected SPI-1 gene expression in growth phases where native tnpA (i.e. in the WT strain) would be limiting relative to invF suggested that native IS200 elements play a role in controlling induction of SPI-1. To characterize the regulatory role of native IS200 elements we compared invF expression in a strain where four of seven IS200 elements were deleted (ΔtnpA2/4/6/7) to the WT strain. RNA was isolated from cells grown to early- and late-exponential phase. In both growth conditions, tnpA expression was reduced ~2.5-fold in the ΔtnpA2/4/6/7 strain, and this correlated with a 2-fold increase in invF expression in early-exponential phase and a 1.5-fold increase in invF expression in late-exponential phase (Figure 7A). The smaller effect of reduced tnpA expression on invF in late-exponential phase is consistent with the above results where invF is present at an excess to tnpA in this growth phase.

We then created a full IS200 knockout strain (ΔtnpA) where all seven copies were deleted. The ΔtnpA strain has a marked growth defect that was suppressed by introducing the tnpA over-expression allele into the ΔtnpA strain (ΔtnpA/tnpA::kan-pTet) (Figure 7B). The complementation strain (ΔtnpA/tnpA::kan-pTet) expresses tnpA at a level much higher than the WT strain (Figure 7D) which suggests that a relatively low amount of tnpA is required for maximal growth. The growth defect in the ΔtnpA strain provides direct evidence that native IS200 elements contribute to host fitness in S. Typhimurium. We next measured invF expression in WT, ΔtnpA and ΔtnpA/tnpA::kan-pTet strains grown to early-exponential phase (OD600 = 0.5). At this growth phase invF was barely detectable by primer extension in the WT strain, but increased dramatically in the strain without IS200 elements (Figure 7C). Importantly, invF expression was reduced to close to WT levels in the complementation strain. Unexpectedly, a second strain encoding tnpA starting at nt 19 (tnpA::kan-pTet (+19)) also suppressed invF expression. This could be an indication that residues in tnpA downstream of the ‘expected’ seed region (see data in Figure 5) play an important role in invF pairing.

To quantitate the effect of native tnpA expression on SPI-1 gene expression we performed RT-qPCR on the RNA used in Figure 7C and D to measure invF, sicA, sipC and prgH expression. Transcript levels of all four genes increased 20- to 25-fold in the ΔtnpA strain and complementation with both of the tnpA::kan-pTet complementation alleles reduced expression close to WT levels (Figure 7E). Surprisingly, prgH expression was also upregulated in the ΔtnpA strain. As PrgH, a component of the T3SS needle complex, is not regulated by invF, this unexpected result could be indicative of tnpA also affecting the expression of a gene upstream of invF in the SPI-1 expression cascade. Lastly, we measured the impact of deleting IS200 elements on the rate of invasion into cultured HeLa cells. Consistent with our expression data, the ΔtnpA strain was 6.7-fold more invasive than the WT strain (Figure 7F).

Overall, the finding that over-expressing and deleting tnpA have opposing effects on invF, sicA and sipC gene expression and Salmonella invasion strongly supports the conclusion that tnpA plays an important role in regulating SPI-1 functions.

DISCUSSION

In the current work we asked if IS200 encoded transcripts affect gene expression in S. Typhimurium. IS200 is an unusual transposon in that it is often present in high copy number in many Salmonella and Yersinia spp. but the transposon itself is almost completely dormant. The low transposition frequency of IS200 can be explained by close to no synthesis of the TnpA protein (1). However, the IS200 transposase mRNA (tnpA) is expressed at a moderate level in S. Typhimurium, resulting in a paradox where this transposon has evolved to maintain transcription of the trans-
posase mRNA but essentially no translation of the protein. Here we provide an explanation for this paradox by demonstrating that over-expression of \( tnpA \) alters the expression of at least 73 genes in \( S. \) Typhimurium, including many genes involved in pathogenesis. We provide evidence that \( tnpA \) is processed to produce small regulatory RNAs that inhibit expression of the SPI-1 encoded transcription factor \( invF \) by a base-pairing mechanism and this impacts on the ability of \( Salmonella \) to invade HeLa cells in vitro.

Ribonucleolytic processing of \( tnpA \) mRNA generates sRNA regulators of \( invF \) expression

We began the current study by profiling the effect of \( tnpA \) over-expression on gene expression in \( S. \) Typhimurium. In this experiment we observed strong repression (>2-fold) of 73 genes, 8 of which (\( sipC, sipA, sseA, sseL, sigE, sopB, sicA, sipB \)) are involved in pathogenesis. Although \( tnpA \) over-expression also represses art200 expression, four of these virulence genes (\( sicA, sipB, sipC \) and \( sopB \)) were repressed by a \( tnpA \) mutant that is unable to downregulate art200. As \( tnpA \) is almost never translated, we speculated that all or part of \( tnpA \) may act as a non-coding RNA to regulate gene expression in \( S. \) Typhimurium.
Figure 7. Contribution of native IS200 elements to regulation of SPI-1 expression. (A) RT-qPCR was performed on WT SL1344 and a derivative in which four of seven copies of IS200 were deleted (ΔtnpA2/4/6/7). RNA was isolated from cells grown to early- or late-exponential phase. Expression of each gene was normalized to the WT strain grown to early-exponential phase. Error bars show the standard error on the mean (n = 4). (B) Growth of the indicated strains in Lennox Broth (LB) was measured in a 96-well microplate spectrophotometer. Error bars (n = 2) are omitted for clarity. Note that the A600 was not adjusted for path length and light scattering of the microplate lid and cannot therefore be directly compared to OD600 measurements of culture density in standard cuvettes. (C–E) RNA was isolated from WT SL1344, a derivative where all seven copies of IS200 were deleted (ΔtnpA), the ΔtnpA strain complemented with over-expression of one copy of tnpA (ΔtnpA/tnpA_7::kan-pTet), and the ΔtnpA strain complemented with over-expression of a 5′ truncated tnpA (ΔtnpA/tnpA_7::kan-pTet(+19)) grown to early-exponential phase (OD600 = 0.5). (C) Primer extension was performed to directly measure invF expression. (D) Northern blot analysis to measure tnpA, art200 and 5S rRNA levels. PCR products were run on the gel (L) to estimate the size of processed tnpA species (indicated with open circles), and the high molecular weight species (closed circle) presumed to be full-length tnpA. (E) RT-qPCR was performed to quantitate transcript levels of invF, sicA, sipC and prgH. Error bars show standard error on the mean (n = 3). (F) HeLa cells were infected with WT or ΔtnpA strains of Salmonella Typhimurium SL1344 (MOI of 100) grown to early-exponential phase (OD600 ~ 0.5). A non-invasive strain (ΔinvA) was also assayed as a control. Bars represent the average invasion for 6 biological replicates (measured in technical duplicate) from two independent experiments. In each experiment, the mean invasion of the WT strain (0.08 and 0.04% of input) was set to 100.
It is now clear that untranslated regions of mRNAs serve as a rich reservoir of sRNAs. As we had observed an effect from over-expressing the 5′ portion of tnpA, we asked if IS200 expresses a 5′ UTR derived sRNA. The typical 5′ UTR derived sRNA (5′ sRNA) is transcribed from the same promoter as an mRNA and transcription terminates at an intrinsic terminator upstream of the coding sequence for the mRNA (12,40,41). Although most 5′ sRNAs terminate at an intrinsic terminator, post-transcriptional processing occurs for several previously described 5′ sRNAs (41–43). Indeed, our primer extension and northern analysis revealed that the 5′ end of tnpA contains two processing sites which produce ∼110-nt RNA initiating at the tnpA transcription start site and ending at nt ∼108 (tnpA-110) and ∼90-nt species (tnpA-90) that is likely generated by processing at nt 19 of tnpA-110. Similar to Type II 3′ UTR derived sRNAs, the tnpA sRNAs are likely stable processing intermediates of tnpA, whereby the biogenesis of the tnpA sRNAs comes as a consequence of ribonucleolytic degradation of an mRNA (35). Evidence for the instability of tnpA 3′ of processing sites comes from the relatively small amount of these downstream products detected by primer extension. At this point we have not identified the ribonuclease responsible for ribonucleolytic processing of tnpA but we predict that RNase III and/or RNase E would be involved based on sequence and structural elements at both processing sites. Future work will investigate the precise mechanism of endoribonucleolytic processing of tnpA including the potential involvement of art200 in generating tnpA-110 and -90.

We found that only the first 50 nt of tnpA is required for repressing InvF and sicAspBC expression which fits fully with tnpA-110 and -90 as acting as a trans-acting sRNA. Mutations to nt 12–19 of tnpA (tnpA119) prevent pairing with InvF in vitro and repression of InvF in vivo, consistent with the first 19 nt containing base-pairing residues. However, as a construct lacking the first 19 nt of tnpA retained the capacity to repress InvF expression, it is likely that residues downstream of nt 19 also contribute significantly to InvF pairing. Consistent with this is the observation that the tnpA Ls mutant impaired tnpA–InvF pairing in vitro (Figure 5).

While we do not yet know how tnpA–InvF pairing represses InvF expression we speculate that pairing primarily leads to degradation of InvF mRNA. The putative tnpA–InvF interaction occurs at the extreme 5′ end of InvF, ∼110 nt upstream of the translation start codon (Figure 5). Base-pairing might recruit RNases to actively degrade InvF mRNA (44). This model is particularly appealing for tnpA-90, as the 5′ monophosphate on this RNA species could directly stimulate RNase E cleavage (45). However, the tnpA complementation allele initiating at nt 19 (tnpA 7::kan- pTet(+19)) would have a 5′ triphosphate and this strain repressed SPI-1 expression to almost the same extent as the full-length tnpA (Figure 7E). Alternatively, either tnpA-90 or -110 may interfere with translation. sRNAs can interfere with translation initiation by base-pairing 50–100 nt upstream of the translation start codon (46–49). The fact that tnpA inhibited InvF protein expression at an early growth phase (Figure 4E) while not affecting InvF mRNA levels (Figure 6E) suggests that translation inhibition is at least one consequence of tnpA–InvF pairing. More work is required to determine the molecular mechanism(s) for how tnpA inhibits invF expression.

While the current work presents the first example of a bacterial transposon producing trans-acting sRNAs, there are two recent examples of transposase derived sRNAs in archaea. The Sulfolobus solfataricus sRNA RNA-257 shares substantial homology with the 3′ UTR of the ISC1904 transposase, ORF1182. RNA-257 is believed to be a remnant of transposition reactions, and this sRNA base-pairs with ORF1183, which encodes a putative phosphate transporter. Similar to tnpA–InvF, base pairing between RNA-257 and ORF1183 results in degradation of the mRNA (8). In Halobacterium salinarum, the IS1341 transposase, tnpB, expresses more than 10 different sRNAs, one of which regulates growth rate by an undetermined mechanism (9). Notably, our observation that deletion of all of the IS200 copies in S. Typhimurium impacted on the expression of an SPI-1 gene (pregH) not under the control of InvF is consistent with tnpA producing either a multi-functional regulatory RNA or, as in the case of IS1341, multiple regulatory sRNAs. Additional studies on the processing of tnpA will be required to further address these possibilities.

It is perhaps surprising that neither tnpA-110 or -90 have been detected in previous work identifying sRNAs in S. Typhimurium (14,24,39). However, a standard practice in mapping RNA-Seq reads to the reference genome is to omit non-unique reads, and the presence of seven identical copies of IS200 in SL1344 would result in reads derived from tnpA being overlooked. However, we note that tnpA is enriched up to 4.1-fold in Hfq-CoIP experiments (14), and the previously characterized Hfq-binding site on tnpA (nt 68–83; (1)) would be present in both tnpA-110 and -90. We have not yet investigated the role of Hfq in tnpA–InvF pairing, in part due to the complications of dysregulated SPI-1 expression (destabilized hiID) (50) in an hfg-null strain.

Regulatory cross-talk between horizontally acquired genes and the S. Typhimurium core genome

*Salmonella* Typhimurium contains a mosaic genome consisting of a core genome complemented with a number of horizontally acquired genetic elements. The core genome is highly conserved among Enterobacteriaceae and contains all of the genes required for normal cellular processes. The accessory or ‘flexible’ genome is made up of a number of horizontally acquired genes including pathogenicity islands, prophage, plasmids and transposons. This flexible genome has been acquired over evolutionary time and provides most of the genes required for virulence (51). Horizontally acquired genes become integrated into host regulatory networks whereby components of the core genome regulate horizontally acquired genes and the core genome itself can be regulated by members of the accessory genome (52). In *S.* Typhimurium, the core genome encoded sRNA SgrS represses expression of the SPI-1 effector *sopD* (53), while the SPI-1 encoded sRNA InvR represses expression of the core genome encoded *ompD* (54). In enterohaemorrhagic *E. coli*, a bacteriophage encoded sRNA, AgvB, represses the core genome encoded sRNA GcvB, thereby increasing expression of many genes involved in amino acid transport (55). In the current work we provide evidence of cross-talk between
members of the accessory genome, where a transposon derived sRNA controls expression of part of the SPI-1 T3SS by repressing expression of invF.

Activation of SPI-1 is controlled by a complex network of transcriptional, post-transcriptional and post-translational regulation (56,57). Environmental signals including low oxygen and high osmolarity first converge to activate transcription of hilD, which in turn activates transcription of structural components of the needle complex (prg/org) as well as invF, a transcriptional activator of SPI-1 effector proteins (sic/sip and other genes). Processed forms of tnpA (tnpA-110 and -90) base-pair with invF mRNA and inhibit InvF expression. Three activation checkpoints for SPI-1 gene expression are indicated with dashed boxes.

Expression of virulence genes has an extreme fitness cost for many bacteria including S. Typhimurium, Y. pestis and Shigella flexneri (61–63). For example, single cell analyses revealed that SPI-1 induction dramatically retarded growth and this growth defect was abrogated by deleting the sic/sip locus (63). In addition, we have shown here that overexpression of invF, sicA and sipC correlates with a reduced growth rate for S. Typhimurium in rich media. As the sicA promoter has the longest relaxation time for SPI-1 encoded genes (64), induction of sic/sip by InvF represents a key commitment step to virulence and the associated burden of producing effector proteins. We propose a model where repression of invF by tnpA sets a threshold for invF induction that must be passed to induce expression of virulence factors (Figure 8). Evidence for this comes from two key experiments. First, when we profiled SPI-1 expression over growth we noted that (i) tnpA was expressed higher than invF in lag and early-exponential phase, and (ii) overexpression of tnpA did not affect invF expression in these two growth phases (Figure 6). Second, reducing tnpA expression 2-fold in early-exponential phase increased invF expression >2-fold, while the same reduction in tnpA expression in late-exponential phase resulted in only a 1.4-fold increase in invF (Figure 7). This threshold for activation would ensure that InvF is only synthesized once there is a sufficiently high transcriptional activation of the invF promoter by HilA. A similar threshold for activation occurs for activation of both hilD and hilA (64,65). In the case of hilD, post-translational repression by HilE dampens hilD activation, and H-NS repression of hilA counteracts transcriptional activation by HilD, HilC and RtsA (66). An additional role of tnpA may be to prevent leaky expression of invF, particularly HilC-dependent activation of the alternative promoter for invF (pinvF-2, Figure 8) (67,68). While hilD, rtsA and hilA are strongly repressed prior to induction, hilC is expressed at a basal level and expression has minor fluctuations independent of hilD (63). As the predicted tnpA–invF base-pairing interaction extends 30 nt upstream...
of the HilA-dependent TSS for invF, the HilC-dependent invF transcript may be subject to even stronger repression by npmA.

Our model predicts that the absence of IS200 elements would lead to premature activation of SPI-1. In a WT strain, invF is induced 6-fold in late-exponential phase and 60-fold in early-stationary phase, when compared to early-exponential phase (Figure 6). In the absence of IS200 elements, invF is induced 25-fold in early-exponential phase when compared to the WT strain (Figure 7E). The early activation of invF expression in the ΔnpmA strain correlates with a 6.7-fold increase in invasion, as well as substantially reduced growth rate (Figure 7). Together, these experiments provide evidence that IS200-encoded RNAs play an important role in delaying the activation of the SPI-1 TSS and that this delayed activation provides a selective growth advantage to S. Typhimurium.

Bacterial transposons as a source of regulatory RNA

Our initial goal for the current work was to determine if any IS200 encoded RNAs affect gene expression in S. Typhimurium. Our transcriptomics experiment identified at least 73 genes that are dysregulated by npmA overexpression. We have investigated how npmA impacts on expression of three of these genes (sicA, sipB and sipC) and our data is consistent with an indirect mechanism where npmA acts through invF to control sicAsipBC expression. We believe that the effect of npmA on many of the other genes identified here will likewise be indirect, and mediated through a smaller number of direct targets. Regardless of the mechanism by which npmA regulates gene expression in S. Typhimurium, we have identified a new way that bacterial transposons can ensure survival: contributing a regulatory RNA. The dual use of a promoter for mRNA and sRNA would ensure that transcription of the transposase is maintained, and post-transcriptional regulation of TnpA expression protects against detrimental effects of transposition.

It is now clear that bacterial sRNAs can be derived from unexpected places. Transposons likely represent an unexplored reservoir of regulatory RNAs that could ultimately provide a benefit to the host organism.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank T. Bloskie for assistance with initial validation of our RNA-Seq data, D. Carter and the staff at the London Regional Genomics Centre for library preparation and Illumina Sequencing, and G. Gloor and J. Macklaim for advice on analyzing RNA-Seq data. We are grateful to M. Valvano for providing S. Typhimurium LT2, J. Vogel for providing the InvF-3xFLAG strain, W. Navarre for providing P22 bacteriophage and B. Coombers for providing the ΔinvA strain. We are also grateful to E. Chau and B. Coombers for advice on invasion assays, and S. Zukowski for culturing HeLa cells.

FUNDING

Natural Sciences and Engineering Research Council of Canada; Canadian Institutes of Health Research (CIHR Grant No. 11281) (to D.H.); Natural Sciences and Engineering Research Council of Canada (NSERC); Schulich School of Medicine and Dentistry; Ontario Graduate Scholarship (to M.E.); Alexander Graham Bell Canada Graduate Scholarship (NSERC CGS-D) (to M.E.). Funding for open access charge: NSERC Discovery Grant. Conflict of interest statement. None declared.

REFERENCES

1. Ellis, M.J., Trussler, R.S. and Haniford, D.B. (2015) A cis-encoded sRNA, Hfq and mRNA secondary structure act independently to suppress IS200 transposition. Nucleic Acids Res., 43, 6511–6527.
2. Beuzon, C.R., Chessa, D. and Casadesus, J. (2004) IS200: an old and still bacterial transposon. Int. Microbiol., 7, 3–12.
3. Signier, P., Perochon, J., Lestradé, L., Mahillon, J. and Chandler, M. (2006) ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res., 34, D32–D36.
4. Signier, P., Gourbeyre, E. and Chandler, M. (2014) Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol. Rev., 38, 865–891.
5. Davies, J. (1994) Inactivation of antibiotics and the dissemination of resistance genes. Science, 264, 375–382.
6. Feschotte, C. (2008) Transposable elements and the evolution of resistance. Nat. Rev. Genet., 9, 397–405.
7. Vollf, J.N. (2006) Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. Bioessays, 28, 913–922.
8. Martens, B., Manoharadas, S., Hasenohrl, D., Manica, A. and Blasi, U. (2013) Antisense regulation by transposon-derived RNAs in the hyperthermophilic archaeon Sulfolobus solfataricus. EMBO Rep., 14, 527–533.
9. Gomes-Filho, J.V., Zaramela, L.S., Italiani, V.C., Baliga, N.S., Vencio, R.Z. and Koshe, T. (2015) Sense overlapping transcripts in IS1341-type transposase genes are functional non-coding RNAs in archaea. RNA Biol., 12, 490–500.
10. Wagner, E.G. and Romby, P. (2015) Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. Adv. Genet., 90, 133–208.
11. Storz, G. and Vogel, J. and Wassarman, K. M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. Mol. Cell, 43, 880–891.
12. Hershko-Shalev, T.,odenheimer-Bergman, A., Elgrably-Weiss, M., Ben-Zvi, T., Govindarajan, S., Seri, H., Papenfort, K., Vogel, J. and Altuvia, S. (2016) Gifsy-1 prophage IsrK with dual function as small and messenger RNA modulates vital bacterial machineries. PLoS Genet., 12, e1005975.
13. Chao, Y. and Vogel, J. (2016) A 3’ UTR-Derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. Mol. Cell, 61, 352–363.
14. 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.
15. Guo, M.S., Updegrove, T.B., Gogol, E.B., Shabalina, S.A., Gross, C.A. and Storz, G. (2014) MicL, a new e-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes Dev., 28, 1620–1634.
16. Jørgensen, M.G., Thomason, M.K., Havelund, J., Valentin-Hansen, P. and Storz, G. (2013) Dual function of the McsA small RNA in controlling biofilm formation. Genes Dev., 27, 1132–1145.
17. Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., Backofen, R. and Vogel, J. (2016) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. EMBO J., 35, 991–1011.
18. van Nues, R.W., Castro-Roa, D., Yuzenкова, Y. and Zenkin, N. (2016) Ribonucleoprotein particles of bacterial small non-coding RNA IsrA (IS61 or McsA) and its interaction with RNA polymerase core may link transcription to mRNA fate. Nucleic Acids Res., 44, 2577–2592.
61. Ali, S.S., Soo, J., Rao, C., Leung, A.S., Ngai, D.H.-M., Ensminger, A.W. and Navarre, W.W. (2014) Silencing by H-NS Potentiates the Evolution of *Salmonella*. *PLoS Pathog.*, **10**, e1004500.

62. Schuch, R. and Maurelli, A.T. (1997) Virulence plasmid instability in *Shigella flexneri* 2a is induced by virulence gene expression. *Infect. Immun.*, **65**, 3686–3692.

63. Sturm, A., Heinemann, M., Arnoldini, M., Benecke, A., Ackermann, M., Benz, M., Dormann, J. and Hardt, W.D. (2011) The cost of virulence: retarded growth of *Salmonella Typhimurium* cells expressing type III secretion system 1. *PLoS Pathog.*, **7**, e1002143.

64. Temme, K., Salis, H., Tullman-Ercek, D., Levskaya, A., Hong, S.-H. and Voigt, C.A. (2008) Induction and relaxation dynamics of the regulatory network controlling the type III secretion system encoded within salmonella pathogenicity island 1. *J. Mol. Biol.*, **377**, 47–61.

65. Saini, S., Ellermeier, J.R., Slauch, J.M. and Rao, C.V. (2010) The role of coupled positive feedback in the expression of the SPI1 type three secretion system in *Salmonella*. *PLoS Pathog.*, **6**, e1001025.

66. Olekhnovich, I.N. and Kadner, R.J. (2006) Crucial roles of both flanking sequences in silencing of the hilA promoter in *Salmonella enterica*. *J. Mol. Biol.*, **357**, 373–386.

67. Lim, S., Lee, B., Kim, M., Kim, D., Yoon, H., Yong, K., Kang, D.-H. and Ryu, S. (2012) Analysis of HilC/D-dependent invF promoter expression under different culture conditions. *Microb. Pathog.*, **52**, 359–366.

68. Akbar, S., Schechter, L.M., Lostroh, C.P. and Lee, C.A. (2003) *AraC/XylS* family members, HilD and HilC, directly activate virulence gene expression independently of HilA in *Salmonella typhimurium*. *Mol. Microbiol.*, **47**, 715–728.