

An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs

Yanjie Chao¹, Kai Papenfort¹,
Richard Reinhardt², Cynthia M Sharma³
and Jörg Vogel^{1,*}

¹Institute for Molecular Infection Biology (IMIB), University of Würzburg, Würzburg, Germany, ²Max Planck Genome Centre Cologne, Max Planck Institute for Plant Breeding Research, Cologne, Germany and ³Research Centre for Infection Diseases (ZINF), University of Würzburg, Würzburg, Germany

The small RNAs associated with the protein Hfq constitute one of the largest classes of post-transcriptional regulators known to date. Most previously investigated members of this class are encoded by conserved free-standing genes. Here, deep sequencing of Hfq-bound transcripts from multiple stages of growth of *Salmonella typhimurium* revealed a plethora of new small RNA species from within mRNA loci, including DapZ, which overlaps with the 3' region of the biosynthetic gene, *dapB*. Synthesis of the DapZ small RNA is independent of DapB protein synthesis, and is controlled by HlID, the master regulator of *Salmonella* invasion genes. DapZ carries a short G/U-rich domain similar to that of the globally acting GcvB small RNA, and uses GcvB-like seed pairing to repress translation of the major ABC transporters, DppA and OppA. This exemplifies double functional output from an mRNA locus by the production of both a protein and an Hfq-dependent *trans*-acting RNA. Our atlas of Hfq targets suggests that the 3' regions of mRNA genes constitute a rich reservoir that provides the Hfq network with new regulatory small RNAs.

The EMBO Journal (2012) 31, 4005–4019. doi:10.1038/emboj.2012.229; Published online 24 August 2012

Subject Categories: RNA

Keywords: ABC transporter; DapZ; GcvB; Hfq; 3' UTR

Introduction

The small non-coding RNAs (sRNAs) that associate with the bacterial RNA-binding protein Hfq have over the recent years emerged as one of the largest class of post-transcriptional regulators (Storz *et al.*, 2011). These sRNAs commonly regulate *trans*-encoded mRNAs by short base pairing, and require Hfq for both their own intracellular stability and their efficient annealing to target mRNAs (Vogel and Luisi, 2011). While the majority of sRNAs negatively regulate their targets by translational repression or destabilization of the mRNA, several examples of mRNA activation have been described (Fröhlich and Vogel, 2009).

*Corresponding author. Institute for Molecular Infection Biology, University of Würzburg, Josef-Schneider-Strasse 2, 97080 Würzburg, Germany. Tel.: +49 931 3182 576; Fax: +49 931 3182 578; E-mail: joerg.vogel@uni-wuerzburg.de

Received: 9 May 2012; accepted: 19 July 2012; published online: 24 August 2012

Similarly to eukaryotic microRNAs, Hfq-associated sRNAs commonly regulate multiple targets to extensively modulate gene expression at the post-transcriptional level. For example, the widely conserved GcvB sRNA targets up to ~1% of all mRNAs in the Gram-negative model species *Salmonella typhimurium* and *Escherichia coli* (Sharma *et al.*, 2007; Pulvermacher *et al.*, 2009; Sharma *et al.*, 2011). Many of the GcvB targets encode ABC transporters of small peptides and amino acids, as well as proteins involved in amino-acid biosynthesis pathways. Comparable pervasive control by sRNAs has been reported in other branches of physiology, e.g., iron usage (Masse and Gottesman, 2002), catabolite repression (Beisel and Storz, 2011) and envelope homeostasis (Gogol *et al.*, 2011). Together with the extensive list of validated mRNA targets, the regulatory scope of Hfq-associated sRNAs has begun to rival that of transcription factors, as illustrated by recently identified new functions in physiological circuits as diverse as biofilm formation (Jorgensen *et al.*, 2012; Mika *et al.*, 2012; Thomason *et al.*, 2012), cell surface modulation (Moon and Gottesman, 2009), amino-acid starvation (Corcoran *et al.*, 2012; Holmqvist *et al.*, 2012), sugar import (Maki *et al.*, 2010; Rice and Vanderpool, 2011), quorum sensing behaviour (Sonnleitner *et al.*, 2011; Shao and Bassler, 2012), switch to anaerobic growth (Boysen *et al.*, 2010; Durand and Storz, 2010) or virulence factor expression (Papenfort *et al.*, 2012).

Nonetheless, although Hfq governs one of the most complex post-transcriptional networks known to date, the true number and nature of the involved sRNAs have largely remained elusive, even in the intensely investigated model species *E. coli* and *Salmonella*. Thus far, roughly two-dozen sRNAs from these two species have been studied in more detail. Almost all of them are transcribed from free-standing sRNA genes located in intergenic regions (IGRs) between annotated protein-coding genes. Yet, there are few commonalities among the sRNAs, as they vary dramatically in length (from 50 to 250 nt), sequence and secondary structure. Short stretches of sequence conservation at the 5' end or within an sRNA sequence typically reveal the base-pairing regions for mRNA selection, which in loose analogy to microRNAs, have been referred to as 'seed' regions (Guillier and Gottesman, 2008; Balbontin *et al.*, 2010; Papenfort *et al.*, 2010). Hfq-binding sites in sRNAs are much less conserved; they have traditionally been seen in A/U-rich single-stranded regions next to a stem-loop structure which often coincides with the ρ -independent transcription terminator found at the 3' end of many sRNAs (Vogel and Luisi, 2011). Arguably, the above features are not limited to transcripts from sRNA genes in IGRs, but can also be envisaged in other types of cellular transcripts, including mRNAs. This complexity is increased by the fact that the ρ -independent terminator itself, i.e., the 3' stem-loop with a poly(U) tail, has been implicated as an additional Hfq-binding site (Otaka *et al.*, 2011; Sauer and Weichenrieder, 2011; Ishikawa *et al.*, 2012). Of note, hundreds of mRNA loci also possess ρ -independent terminators

(Lesnik *et al*, 2001; Kingsford *et al*, 2007), and many mRNA 3' regions are highly enriched by co-immunoprecipitation with Hfq (Zhang *et al*, 2003; Sittka *et al*, 2008). Thus, one may speculate that by virtue of binding to Hfq, such transcripts may attain a regulatory function that is independent of the protein encoded by the mRNA of a given locus.

In this study, we have harnessed RNA deep sequencing to reveal a dynamic landscape of Hfq-bound transcripts in *Salmonella* at various stages of growth. We report dramatic changes in the profiles of Hfq-associated sRNAs, including the transient appearance of some sRNAs between the exponential and stationary phases of growth. The profiles reveal many sRNAs from the 3' regions of mRNA loci that are produced by either mRNA processing or overlapping sense transcription with a shared terminator.

One of these novel sRNAs is DapZ, which we have investigated in detail with respect to the mechanisms of its biogenesis and physiological function. This sRNA overlaps in sense with the 3' UTR of the wide-spread biosynthetic *dapB* gene, and is transcribed from a conserved gene-internal promoter which in *Salmonella* evolved to be co-activated with major virulence genes. By employing a G/U-rich seed domain reminiscent of the globally acting GcvB sRNA (Sharma *et al*, 2011), DapZ represses the mRNAs of major ABC transporters under conditions of *Salmonella* host cell invasion. Our results suggest that the 3' regions of mRNA genes constitute a large reservoir from which the Hfq network recruits new sRNAs to rewire gene expression at the post-transcriptional level.

Results

A dynamic landscape of Hfq-associated sRNAs

We profiled Hfq-associated transcripts in *Salmonella* using Solexa sequencing of RNA enriched by co-immunoprecipitation (coIP) with the chromosomally encoded, epitope-tagged Hfq protein (Sittka *et al*, 2008). Samples were collected at several time points along the growth curve of a shaking culture within 1 day of inoculation (Figure 1A), and covered the exponential phase of growth (OD₆₀₀ of 0.15 or 0.5), the early stationary phase (OD₆₀₀ of 2) when the *Salmonella* invasion genes are transiently activated, and the stationary phase at four subsequent time points including overnight culture.

Figure 1B depicts the individual sRNA species enriched by pull-down with Hfq protein at the selected seven points of growth, as a proportion of the experimentally validated sRNAs of *Salmonella* annotated in this (Supplementary Table S1) and previous studies (Padalon-Brauch *et al*, 2008; Sittka *et al*, 2008; Sittka *et al*, 2009; Kröger *et al*, 2012). We observed a dramatic change in the profiles along the axis of growth, with some sRNAs dominating in individual growth phases. For example, the catabolite repression-associated Spot42 (Spf) sRNA prevails in fast growing cells, whereas a different set of sRNAs including RprA, SdsR and RybB become prominent partners of Hfq as bacteria progress into the stationary phase. There are only two abundant sRNAs—ArcZ (activator of σ^S and repressor of serine uptake and oxidative stress related functions) and ChiX (repressor of chitoporin synthesis)—which significantly occupy Hfq throughout growth, and both are well-known examples of growth phase-independent sRNA expression (Argaman *et al*, 2001; Vogel *et al*, 2003; Figueroa-Bossi *et al*, 2009; Papenfort

et al, 2009; Mandin and Gottesman, 2010; Rasmussen *et al*, 2009). For other well-studied sRNAs, Hfq occupancy was also in excellent agreement with previously determined expression profiles in *Salmonella*. For example, GcvB accumulates in fast growing cells only (Argaman *et al*, 2001; Sharma *et al*, 2007), whereas SdsR and RybB are not expressed until stationary phase when their transcription is activated by the alternative σ^S and σ^E factors, respectively (Vogel *et al*, 2003; Papenfort *et al*, 2006; Fröhlich *et al*, 2012).

A distinct set of sRNAs mark the transition from exponential to stationary phase, as exemplified by the invasion gene-associated InvR sRNA of *Salmonella* pathogenicity island 1 (SPI-1). The SPI-1 locus is transiently transcribed as aerobic cultures reach an OD₆₀₀ of 2 (Pfeiffer *et al*, 2007), and it is this cDNA library that contains the vast majority of InvR reads. Other prominent sRNAs in the transition phase are RybD, STnc440 and DapZ; the latter was here renamed from candidate STnc820 due to a genetic association with *dapB* (see below).

Regarding protein-coding transcripts, a total of 3517 mRNAs were detected (represented by ≥ 10 reads) at the different stages of growth, and 1253 of them were enriched at least three-fold by coIP with Hfq (Supplementary Table S2), which expands the putative Hfq regulon to more than a quarter of all *Salmonella* genes. All in all, the new Hfq profiles reveal that the well-studied changes of primary gene expression over growth are accompanied by consistent and dynamic changes of RNA binding to Hfq, the major hub of post-transcriptional control.

Hfq-bound 3' regions of mRNA genes accumulate as discrete sRNAs

While many of the abundantly recovered sRNAs are transcribed from free-standing genes, e.g., ArcZ, ChiX, InvR, RprA, RybB, SdsR and Spot42, we noticed that libraries from all growth phases contained ample cDNA reads that overlapped in sense with the 3' UTR of mRNAs, e.g., DapZ, STnc840, STnc850 and STnc870 (Figures 1B and 2; Supplementary Table S1). Note that for simplicity, we refer to 3' UTR as either the RNA or DNA region that follows the coding sequence of the respective gene down to the transcription terminator. In other words, 3' UTR can denote either the transcribed 3' part of the mRNA or its corresponding DNA in the genome, depending on context. To test whether these cDNAs represent discrete RNA species, rather than premature termination products of cDNA synthesis at the 3' end of mRNA, we selected candidate regions with both a high cDNA count and a predicted ρ -independent terminator for northern blot analysis (Supplementary Table S1).

Of 22 candidates tested, 8 yielded discrete hybridization signals (Figure 2) that agreed well with the respective transcript sizes predicted by Solexa sequencing (Supplementary Table S1). We observed excellent correlations between the northern blot signals of individual sRNAs and their relative coverage in the cDNA libraries. For example, DapZ and STnc840 whose expression sharply peaks at OD₆₀₀ of 2 also show highest recovery in the corresponding cDNA library (compare Figures 2 and 1B, or Supplementary Table S3). Similarly, the northern blots confirmed that STnc850 and STnc870 strongly accumulate in late stationary phase, as predicted by the fact that these sRNAs together constitute 10–25% of all reads in the corresponding cDNA libraries. In several cases, we also detected the mRNA of the same locus,

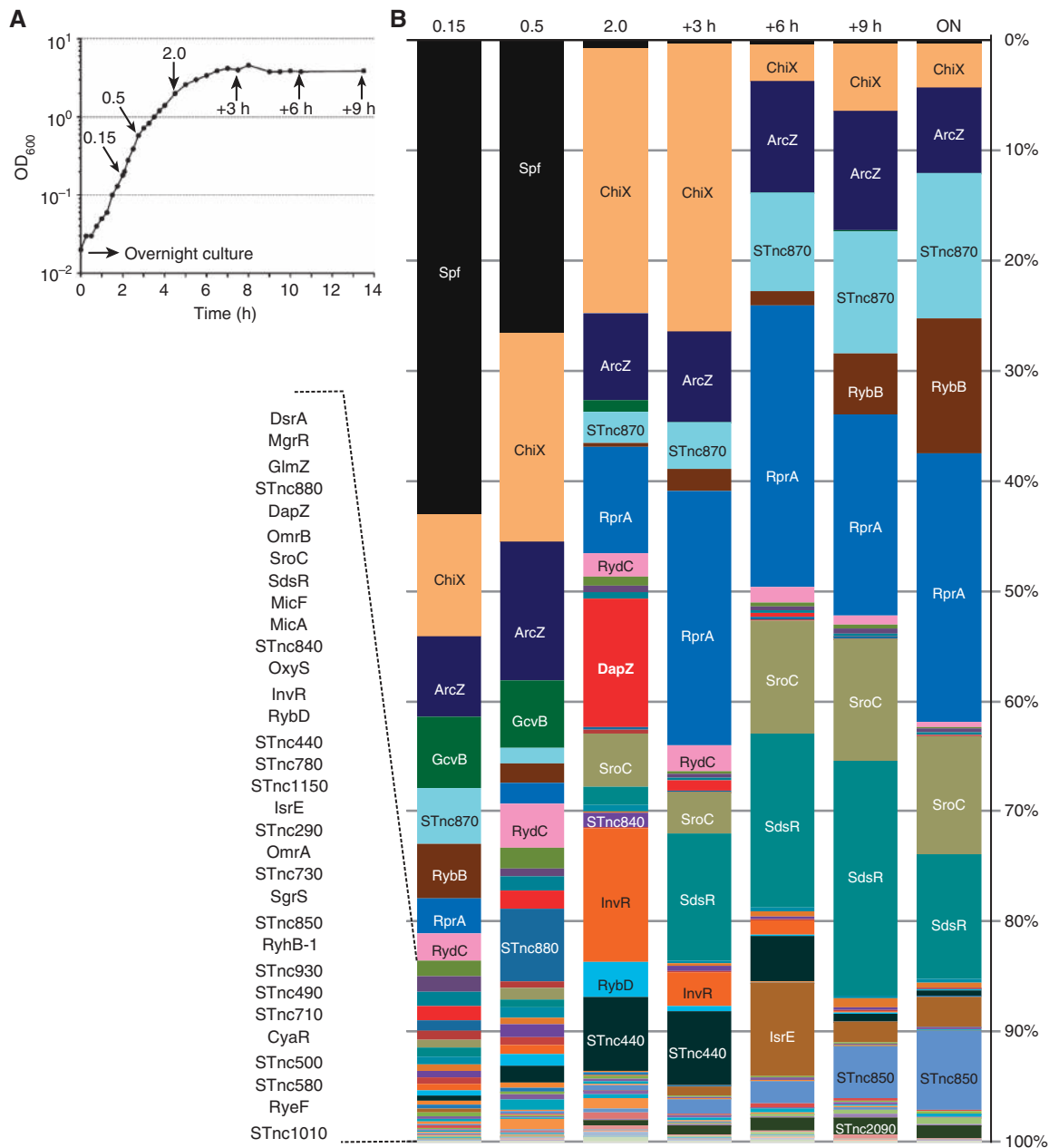


Figure 1 Dynamic sRNA profiles of Hfq over bacterial growth. **(A)** Growth curve of *Salmonella* grown for 14 h in LB at 37°C, 220 r.p.m. Time points when culture samples were withdrawn for Hfq co-immunoprecipitation are indicated. **(B)** Distribution of reads matching experimentally validated sRNAs in Hfq-coIP cDNA libraries at several stages of growth. Percentage indicates the reads of a given sRNA compared to all sRNAs in a cDNA library. The relative amount of reads and enrichment factors for individual sRNAs are listed in Supplementary Table S3. ON: overnight culture.

with expression patterns that either matched (STnc850/*ycfJ* or STnc2090/*yfiA*) or deviated from (DapZ/*dapB* or STnc870/*cpXP*) the sRNA in its 3' region. We consistently observed much stronger sRNA signals than mRNA signals, supporting the notion that these 3'-derived sRNAs may accumulate to fulfill an mRNA-independent function.

An Hfq-dependent sRNA from the 3' region of the *dapB* gene

For proof-of-principle that 3' UTR-derived sRNAs are functional regulators, we focussed on the ~80-nt DapZ sRNA that coincided with the 3' UTR of *dapB* (Figure 2). The *dapB* gene encodes

dihydrodipicolinate reductase, an essential protein that catalyses the second step of lysine biosynthesis (Bouvier *et al*, 2008a). Intriguingly, whereas the coding sequence of *dapB* is highly conserved in γ -proteobacteria (Supplementary Figure S1), the 3' UTR is not, except for the ρ -independent terminator (Figure 3A).

We considered that DapZ may be either a stable intermediate of mRNA decay, as suggested by its mutually exclusive accumulation with the full-length *dapB* transcript (Figure 2; Supplementary Figure S2), or the product of a hidden sRNA gene. Three experimental results confirmed the latter scenario, i.e., that DapZ is an independently transcribed sRNA,

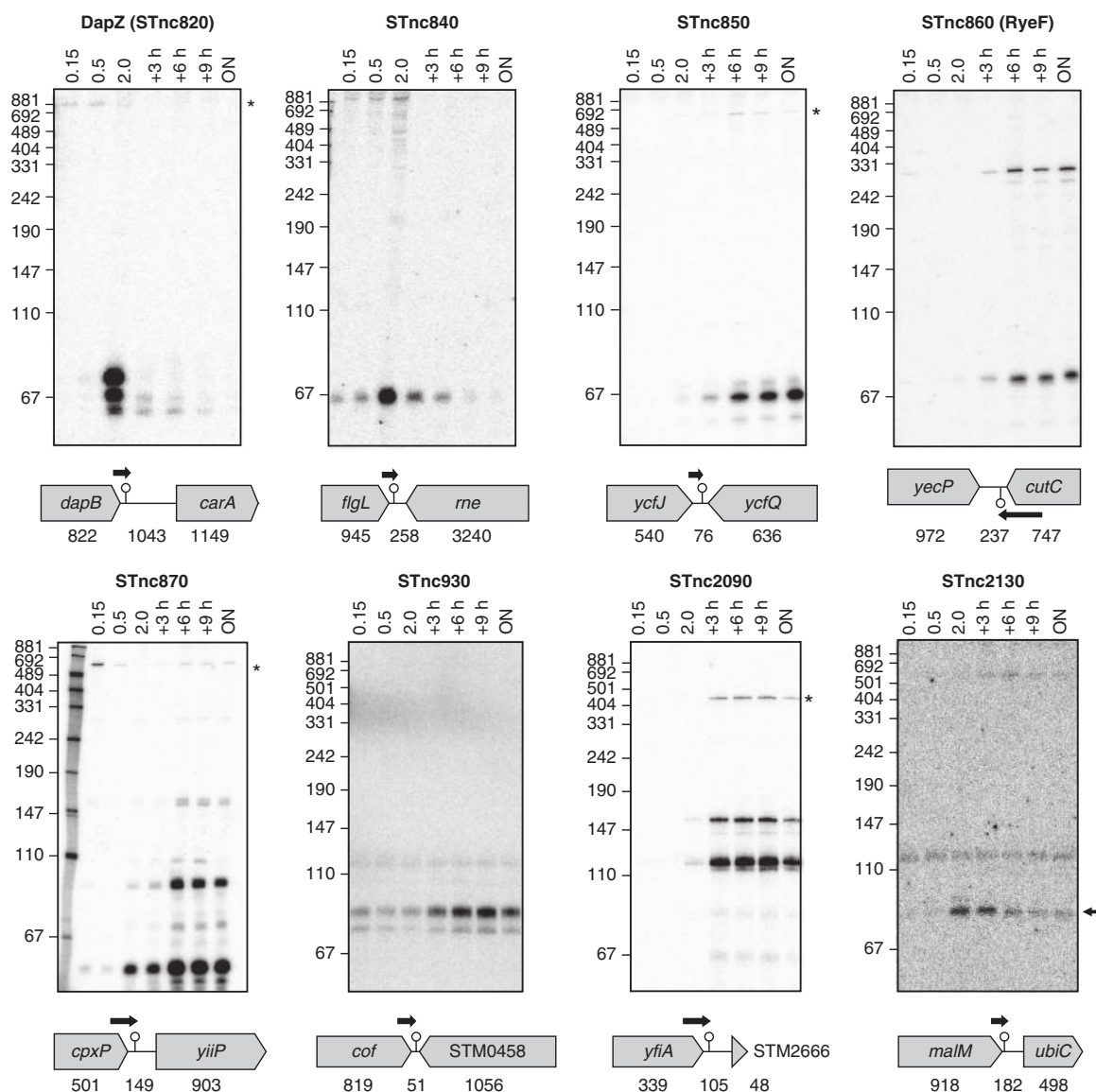


Figure 2 Expression analysis of 3' UTR-derived sRNAs. Total RNA was prepared from wild-type *Salmonella* grown in LB at the time points indicated in Figure 1A, and subjected to northern blot analysis. (*) denotes detection of associated full-length mRNAs in the cases of DapZ, STnc850, STnc870 and STnc2090. The position of sRNAs, the name and the length of flanking genes, as well as the length (bp) of the intergenic regions are shown in the schematic presentations below the blots. All sRNAs identified are in close proximity to, or even partially overlap with upstream genes. Hybridization probes are listed in Supplementary Table S8.

encoded within the 3' region of *dapB*. First, the presence of a 5' truncated, promoterless *dapB* gene on a multi-copy plasmid (*pdapZ*) caused overexpression of DapZ (Figure 3B), indicating that DapZ synthesis did not require transcription of *dapB*. Second, using a specialized 5' RACE protocol we detected a transcriptional start site (TSS) immediately downstream of the *dapB* stop codon that coincided with the 5' flank of reads in the cDNA libraries (Figure 3C), and was preceded by well-conserved putative -10 and -35 boxes in the *dapB* CDS (Figure 3A). Assuming that DapZ transcription terminates at the *dapB* terminator, this TSS perfectly matches the observed sRNA size on the northern blot (Figure 2). Third, the main DapZ transcript is resistant to treatment with a 5' monophosphate-dependent exonuclease (Kröger *et al*, 2012), which is consistent with DapZ being a primary transcript that possesses a 5' triphosphate.

A comparison of RNA abundance and stability between *Salmonella* wild-type and an isogenic Δ *hfq* mutant in early stationary phase confirmed that DapZ is an abundant Hfq-dependent sRNA. At the height of its expression, DapZ accumulates to ~ 100 copies/cell, but the lack of Hfq reduces the *in vivo* copy number to less than one (Supplementary Figure S3). The RNA half-life was reduced from ~ 2.5 min to < 30 s in the Δ *hfq* strain, suggesting that as with many other sRNAs (Andrade *et al*, 2012), Hfq protects DapZ from degradation. Note that we also observed shorter, processed DapZ species in both the RACE and northern blot experiments. However, these partially or fully lack the seed region of DapZ (as defined below) and so are likely non-functional decay intermediates of the sRNA. Altogether, these results established that DapZ is an Hfq-dependent sRNA that shares the 3' region and terminator of the *dapB* mRNA

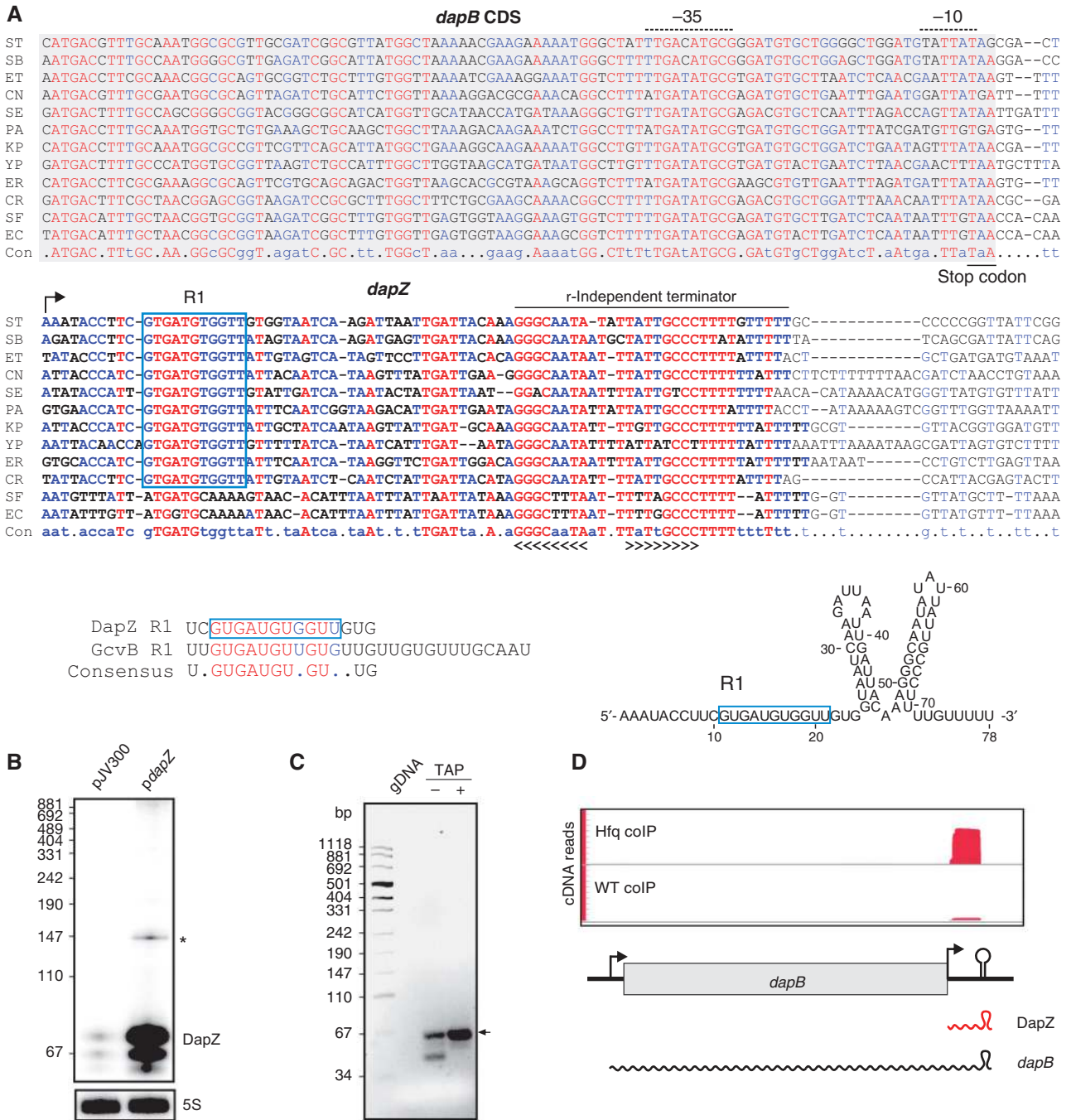


Figure 3 Promoter and sequence analysis of DapZ sRNA. (A) Sequence alignment of the *dapB* 3' coding sequence (CDS, grey box) and 3' UTR of related enterobacterial species. The *Salmonella* *dapZ* sequence and its homologous sequences in other species are shown in bold. The arrow denotes the +1 site of *dapZ* in *Salmonella*. Putative promoter motifs within the *dapB* CDS and the ρ -independent terminator sequence are indicated. The conserved GU-rich motif R1 (boxed) is found in most species but absent in *E. coli* and *Shigella*. ST: *Salmonella typhimurium*; SB: *S. bongori*; ET: *Enterobacter* spp.; CN: *Cronobacter* spp.; SE: *Serratia* spp.; PA: *Pantoea* spp.; KP: *Klebsiella pneumoniae*; YP: *Yersinia pestis*; ER: *Erwinia* spp.; CR: *Citrobacter rodentium*; SF: *Shigella flexneri*; EC: *E. coli*; Con: consensus sequence. Below left: DapZ GU-rich motif R1 (boxed) displays high similarity to GcvB R1. Below right: The secondary structure of DapZ sRNA (the GU-rich motif R1 is boxed) predicted by Mfold and validated by structure probing (Supplementary Figure S11). (B) The DNA sequence downstream of the *dapB* start codon down to the ρ -independent terminator was cloned into a high-copy plasmid, and the expression of DapZ in wild-type *Salmonella* at OD₆₀₀ of 2 was determined by northern blot. (*) denotes read-through to the *rnmB* terminator encoded on the plasmid. (C) Identification of the primary transcription start site of *dapZ* by 5' RACE. PCR products were analysed on a 4% agarose gel. The arrow indicates the band corresponding to the primary transcript, which is enriched by RNA pre-treatment with TAP. The appearance of a weaker RACE signal for the full-length DapZ sRNA in the '– TAP' lane may be attributed to the activity of RppH or a related pyrophosphohydrolases (Deana et al, 2008), which converts to 5'PPP to 5'P ends *in vivo*. The shorter RACE product in the TAP – lane represents a processing intermediate at nucleotide U12 of DapZ (unpublished results), and as expected, is not found in the TAP + reaction. (D) Schematic drawing showing that the DapZ sRNA is transcribed from the 3' UTR of the *dapB* gene and enriched at early stationary phase in the Hfq colIP library.

gene (Figure 3D). Remarkably, both its proximity to the *dapB* reading frame and its low degree of sequence conservation would prevent faithful prediction of this sRNA by biocomputational methods.

The master regulator of virulence HilD regulates *dapZ*

Although their genes partly overlap, the *dapB* mRNA and DapZ sRNA are clearly differentially expressed (Figure 2; Supplementary Figure S2). Expression of *dapB*, which according to *E. coli* data may be regulated by lysine and the LysR-type regulator ArgP (Bouvier *et al*, 2008a), was highest during exponential growth. By contrast, DapZ sharply accumulated in the transition to the stationary phase, matching the profile of the SPI-1-associated InvR sRNA (Figure 1B; Pfeiffer *et al*, 2007), which raised the intriguing possibility that DapZ was co-activated with the SPI-1 invasion genes. This hypothesis was supported by northern blot analysis of *Salmonella* strains with genomic deletions of the major virulence islands. Figure 4A shows that a Δ SPI-1 mutation diminished DapZ expression, whereas a Δ SPI-2 mutation, affecting the genes for intracellular survival, did not affect DapZ expression.

Next, to identify the responsible transcription factor, we investigated DapZ expression in strains with individual disruptions of the *hilA*, *hilC*, *hilD*, *invF* and *rtsAB* genes, which encode transcriptional activators that control the SPI-1 genes in a hierarchical manner (Ellermeier and Slauch, 2007). These experiments suggested HilD, which acts at the very top of the SPI-1 cascade, as the cognate activator of DapZ expression (Figure 4A). The regulation is also observed with a reporter gene fusion to the *dapZ* promoter, arguing that HilD indeed regulates DapZ at the level of transcription (Supplementary Figure S4).

Ectopic expression of HilD restored DapZ levels in *Salmonella* Δ *hilD* or Δ SPI-1 strains (Figure 4B). Again, this was specific to HilD, since the downstream acting HilA factor could not restore DapZ levels in these mutant backgrounds. Most importantly, however, HilD protein also activated the *Salmonella* *dapZ* promoter after transfer to *E. coli*, an organism that lacks SPI-1 and all of its associated transcription factors (Figure 4C). Although a molecular interaction remains to be proven, this *trans*-complementation strongly argues that HilD activates the *dapZ* promoter directly.

The results of additional *trans*-complementation experiments in *E. coli* predict that the control of the widely conserved *dapZ* gene by the *Salmonella*-specific factor HilD evolved very recently, i.e., after *Salmonella* diverged from the *E. coli* lineage. That is, of five selected enterobacterial *dapZ* promoters tested in this assay (Figure 4C), the *Salmonella* promoter was the only one to respond to HilD, but did so as strongly as the *hilA* promoter, a positive control and well-established direct target of HilD (Schechter and Lee, 2001). By contrast, the other four *dapZ* promoters were invariably insensitive to HilD expression, and either showed intermediate (*Citrobacter*, *Klebsiella* and *Yersinia*) or very high basal expression (*E. coli*). Thus, closely related enterobacteria may use different factors to control the *dapZ* promoter.

DapZ sRNA regulates major ABC transporters with a GcvB-like seed domain

Hfq-associated sRNAs typically regulate gene expression by base pairing with mRNAs. To identify targets of DapZ in *Salmonella*, we took a pulse-expression approach (Masse

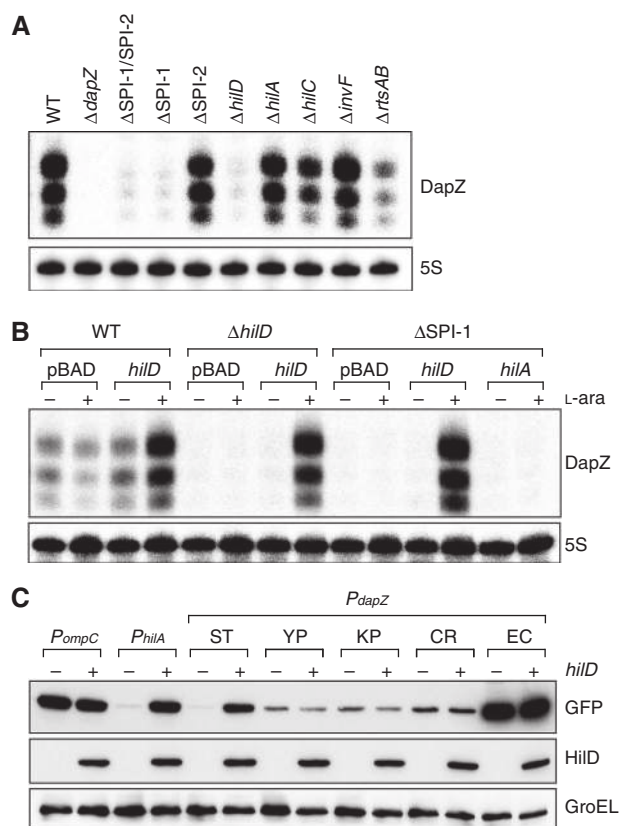


Figure 4 DapZ is activated by the *Salmonella*-specific virulence regulator HilD. (A) *Salmonella* wild-type as well as several deletion mutant strains were grown in LB to an OD₆₀₀ of 2, and total RNA was probed for DapZ expression by northern blot. 5S rRNA served as loading control. (B) *Salmonella* wild-type, Δ *hilD* and Δ SPI-1 strains were transformed with an arabinose-inducible pBAD control plasmid, pBAD-*hilD* or pBAD-*hilA* (Δ SPI-1 strain) and cultivated in LB to OD₆₀₀ of 1.0. Expression from pBAD plasmids was induced by addition of 0.2% L-arabinose (final conc.) for ~45 min and DapZ levels were determined by northern blot. (C) Western blot analysis of GFP reporters in which *dapZ* homologues from several related enterobacteria were fused to a promoterless *gfp* gene: *Salmonella typhimurium* (ST), *Yersinia pestis* (YP), *Klebsiella pneumoniae* (KP), *Citrobacter rodentium* (CR) and *E. coli* (EC). *PompC* and *PhilA* are the control promoter regions known to be non-responsive or responsive, respectively, to HilD. *E. coli* co-transformed with the indicated GFP-reporter plasmids as well as the pBAD-*hilD* (*hilD*) plasmid were grown in LB for 2 h in presence of 0.0004% L-arabinose after reaching OD₆₀₀ of 0.5. Probing for GroEL served as loading control.

et al, 2005; Papenfort *et al*, 2006) analysing global mRNA level changes after a transient overexpression of DapZ from an inducible plasmid. Within 10 min of induction, DapZ downregulated the conserved *dpp* and *opp* operons, which encode major ABC transporters, and the *yahO* and STM1513 genes of unknown function ≥ 2 -fold (Figure 5A). This rapid downregulation suggests that DapZ regulates these mRNAs directly.

Intriguingly, the DapZ candidate targets appeared to be a subset of the many targets of the widely conserved Hfq-associated sRNA, GcvB (Figure 5B). Moreover, the putative DapZ targets overlapped specifically with those mRNAs that GcvB recognizes by its G/U-rich seed domain (Sharma *et al*, 2011), which suggested that DapZ may select targets by a similar mechanism. Indeed, inspection of the DapZ sequence

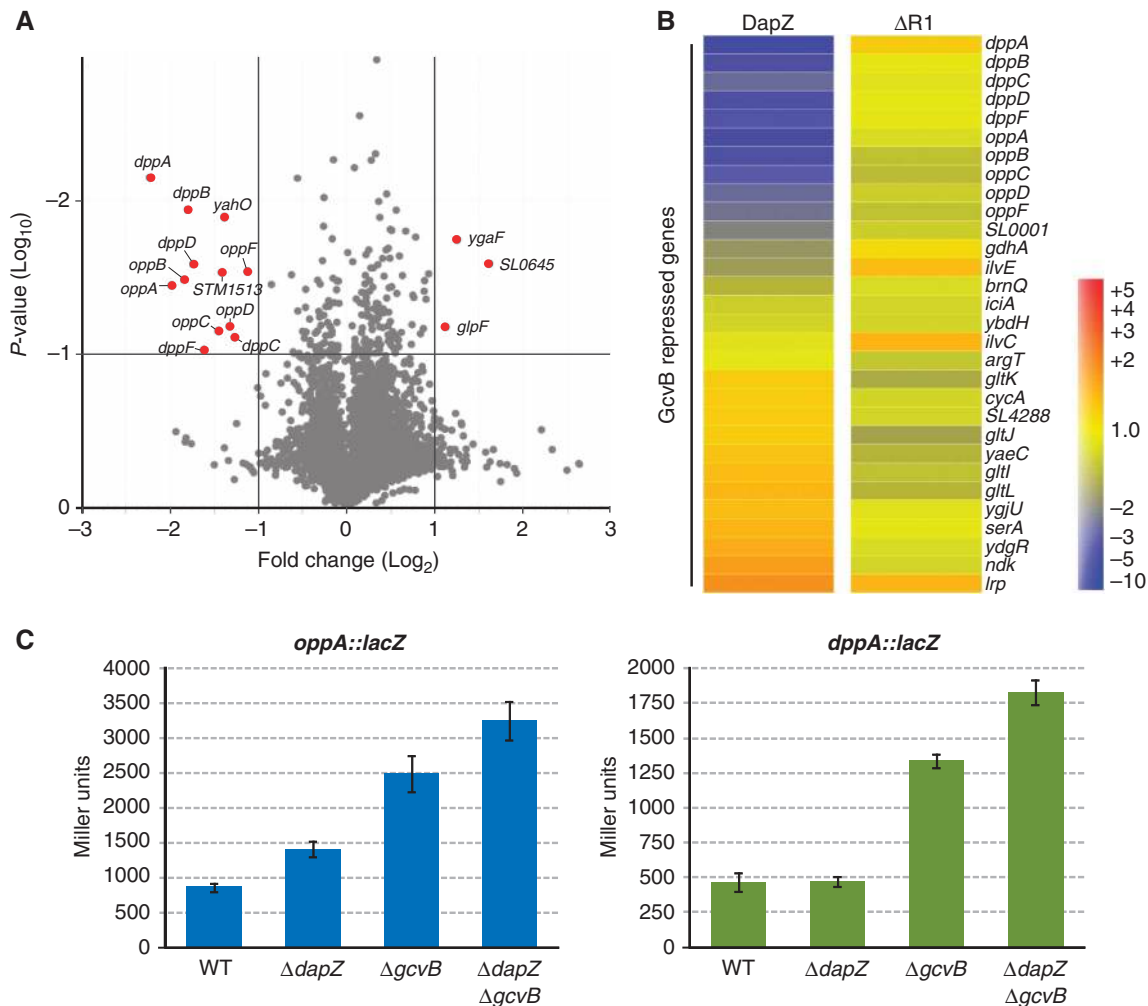


Figure 5 DapZ is a repressor of the *opp* and *dpp* operons. **(A)** Microarray analysis of genes affected by pulse expression of DapZ compared to pBAD control vector in *Salmonella*. *Salmonella dapZ* mutants containing a pBAD control vector or pBAD-DapZ plasmid were grown in LB until OD₆₀₀ of 1.5 and then 0.2% L-arabinose was added to both cultures for 10 min to induce DapZ expression (Supplementary Figure S12). Global transcriptome changes were scored on *Salmonella*-specific microarrays. Genes which show ≥ 2 -fold change (P -value < 0.1) are marked in red. **(B)** Heat map analysis of microarray results of genes regulated by pulse expression of wild-type DapZ and DapZ- $\Delta R1$. All the experimentally validated direct targets of the *Salmonella* GcvB sRNA (Sharma *et al*, 2007, 2011) are shown, and the fold-change values are listed in Supplementary Table S5. **(C)** Endogenous DapZ represses *oppA* and *dppA* protein synthesis in *Salmonella* at OD₆₀₀ of 2. Translational *lacZ* fusions were constructed in the *Salmonella* chromosome by fusing *lacZ* to the 17th codon of *oppA* or the 10th codon of *dppA*. β -Galactosidase activity in wild-type, $\Delta dapZ$, $\Delta gcvB$ and $\Delta dapZ \Delta gcvB$ double mutant *Salmonella* was determined in triplicates.

revealed a single-stranded GUGAUGUGGUU (nucleotides 11–21) stretch that is conserved in enterobacteria except for *E. coli* and *Shigella* (Figure 3A). In analogy to its counterpart in GcvB, this G/U-rich stretch will be referred to as domain R1 of DapZ (Figure 3A). By repeating the pulse-expression experiment with an R1 mutant of DapZ (internal deletion of nucleotides G₁₃A₁₄U₁₅G₁₆), we observed that this domain was indeed essential for the repression of *dpp* and *opp* operons (Figure 5B).

GcvB represses the *dpp* and *opp* operon mRNAs through base-pairing interactions with the first cistrons, *dppA* and *oppA*, respectively (Sharma *et al*, 2007; Pulvermacher *et al*, 2008). Using translational *lacZ* fusions to the chromosomal *dppA* and *oppA* genes of *Salmonella*, we found that DapZ also exerts negative regulation at the 5' end of these operons. A $\Delta dapZ$ mutation activated the *oppA::lacZ* reporter, both alone and further with a $\Delta gcvB$ mutation. The *dppA::lacZ* reporter was also upregulated upon deletion of *dapZ*,

although this required prior inactivation of the *gcvB* gene (Figure 5C). Perhaps when GcvB is present, it occupies the *dppA* target site owing to its more stable binding relative to DapZ (Sharma *et al*, 2007), so the *dppA* target is only bound and suppressed by DapZ when its competitor GcvB RNA is absent. Northern blot analysis showed that the two sRNAs did not influence each other's expression (Supplementary Figure S5), suggesting that DapZ is able to regulate *dppA* and *oppA* independently of GcvB. Likewise, the expression of *dapB* was not significantly affected by the $\Delta dapZ$ mutation (Supplementary Figure S6) or DapZ overexpression (Supplementary Figure S7), ruling out the possibility of indirect effects through an impaired metabolic function of DapB. Altogether, these experiments showed that the major ABC transporters are regulated independently of protein output from the *dapBZ* locus, and that the DapZ sRNA resembled GcvB such that it repressed some shared targets, likely via a similar seed domain.

Evidence for similar seed pairing by DapZ and GcvB

To elucidate how DapZ recognizes the *dppA* and *oppA* targets, we subjected *in vitro* transcribed DapZ and the 5' mRNA regions to structure probing with single strand-specific RNase T1 or lead(II), alone and after mixing sRNA and target. Preliminary gel-shift assays confirmed that DapZ formed complexes with the *oppA* and *dppA* RNA fragments (Supplementary Figure S8). Probing of the 5' labelled DapZ RNA showed that presence of either the *dppA* and *oppA* RNAs protected the G/U-rich R1 domain of the sRNA from cleavage, which was most pronounced with lead(II) backbone cleavage (Figure 6A). Reciprocally, probing of the targets identified DapZ-induced protection of C/A-rich sites around the start codon of *oppA*, or upstream of the Shine-Dalgarno sequence in *dppA*, as well as a structural rearrangement in *dppA* (Figure 6B and C). In other words, although the target interactions of DapZ and GcvB slightly differ with respect to helix length and continuity (Figure 6E), the two sRNAs essentially recognize the same regions in each of the two mRNAs.

To prove that the predicted RNA duplexes guided DapZ-mediated mRNA repression *in vivo*, we introduced compensatory point mutations in DapZ and the *dppA* or *oppA* mRNAs (Figure 6D and E). DapZ was constitutively expressed from a plasmid in a Δ *dapZ* Δ *gcvB* strain, and target regulation was monitored using translational *dppA::gfp* or *oppA::gfp* fusions, respectively, on compatible plasmids (Sharma *et al*, 2007; Urban and Vogel, 2007). Flow cytometry-based measurements of GFP fluorescence showed that DapZ repressed the *oppA::gfp* and *dppA::gfp* reporters 6-fold and 2.3-fold, respectively. Repression was abolished with DapZ-M1, a mutant sRNA with a single G₁₃→C change in the G/U-rich seed (Figure 6D and E). Likewise, M1' variants of the *dppA::gfp* and *oppA::gfp* reporters containing the opposite C→G point mutation were insensitive to wild-type DapZ. However, restoration of the predicted RNA interactions by combining DapZ-M1 with the M1' reporters of *dppA* or *oppA* fully restored repression, which validates that DapZ employs its G/U-rich seed to repress *dppA* and *oppA* by direct base pairing *in vivo*.

Discussion

Expression of a bacterial mRNA gene is generally assumed to culminate in a single functional gene product, which is the protein to be translated from its reading frame. This simple structure–function relationship is regarded as a characteristic feature of bacterial mRNA loci, in contrast with their eukaryotic counterparts which can produce multiple protein isoforms through mRNA splicing as well as non-coding RNAs with independent functions (Rodriguez *et al*, 2004). However, this study suggests that bacterial mRNA loci offer greater functional output than assumed, and furnish evidence for the hypothesis that UTRs can evolve to produce sRNAs that regulate gene expression *in trans* (Vogel *et al*, 2003).

Our deep-sequencing experiments reveal a dynamic re-patterning of Hfq-associated sRNAs at multiple stages of growth, and the single-nucleotide resolution of the technique allowed us to precisely map several novel Hfq-bound species derived from the 3' regions of mRNA genes. We have shown that the 3' UTR-encoded DapZ sRNA is a *trans*-acting regulator with a GcvB-like seed domain that is transcriptionally

activated by the horizontally acquired transcription factor HldD to repress ABC transporter synthesis under conditions that favour host cell invasion. The discovery of DapZ argues that transiently accumulating sRNA species from the 3' end of mRNA loci must not be generally dismissed as noise resulting from spurious transcription or incomplete transcript degradation. Since the 3' UTR is a genomic element that is present in most if not all bacteria (Kingsford *et al*, 2007), it should be systematically explored for new regulatory small RNAs.

Discovery of sRNAs via Hfq profiling

Biocomputational searches have traditionally focussed on the discovery of free-standing sRNA genes in the IGRs of bacterial chromosomes, and relied much on the conservation of transcription elements including the 3' terminal ρ -independent terminators (Vogel and Sharma, 2005; Backofen and Hess, 2010). However, 3' UTR-derived sRNAs such as DapZ pose a challenge to *in-silico* prediction owing to their close proximity or even overlap with mRNA sequences. The *dapZ* gene indeed escaped detection in biocomputational sRNA screens (Argaman *et al*, 2001; Rivas and Eddy, 2001; Wassarman *et al*, 2001; Chen *et al*, 2002; Livny *et al*, 2006; Pfeiffer *et al*, 2007), which we attribute to both its poorly conserved primary sequence and proximity to the *dapB* reading frame. Interestingly, DapZ was also not recognized in cDNA cloning-based screens (Vogel *et al*, 2003; Kawano *et al*, 2005) or earlier analyses of Hfq-bound transcripts by tiling arrays (Zhang *et al*, 2003) or deep sequencing (Sittka *et al*, 2008). By contrast, our present study builds upon deep profiling of Hfq ligands over growth and dedicated inspection of UTR-derived transcripts (Figure 1B). In addition, since Hfq is a limiting factor *in vivo* (Lease and Woodson, 2004; Fender *et al*, 2010; Moon and Gottesman, 2011), the high recovery of DapZ (~12% of reads in the transition phase library) was a strong predictor of a physiological function. That is, other sRNAs that dominate the Hfq profile at this condition are known to be functional: ChiX and InvR repress porin synthesis (Pfeiffer *et al*, 2007; Figueroa-Bossi *et al*, 2009; Rasmussen *et al*, 2009), and ArcZ and RprA regulate the *rpoS* mRNA and additional targets (Papenfort *et al*, 2009; Mandin and Gottesman, 2010). By the same token, prominent Hfq-binding sRNAs from 3' UTRs which accumulate in different growth phases may similarly turn out as bona fide regulators of gene expression.

The general trends of sRNA expression are well reflected by our Hfq profiling over growth, as clearly seen for the abundant marker sRNAs such as GcvB, DapZ/InvR and RprA/SdsR/RybB, which accumulate in early, middle or late growth stages, respectively (Figure 1B). Copy numbers are available for several *Salmonella* sRNAs, and often match the relative recovery of a given sRNA. For instance, both DapZ and InvR accumulate to ~100 copies/cell in early stationary phase (Supplementary Figure S3; Pfeiffer *et al*, 2007), and are equally represented in the corresponding library (Figure 1B). Under the same condition, ArcZ and SdsR are present in ~20 copies/cell, respectively (Papenfort *et al*, 2009; Fröhlich *et al*, 2012), and irrespective of deviations from expected numbers, these sRNAs generally exhibit lower recovery than DapZ and InvR (Figure 1B). Nonetheless, many factors may influence the recovery rate, ranging from non-linear binding to Hfq to sRNA-specific biases during cDNA preparation and sequencing. Thus, low abundance cannot be

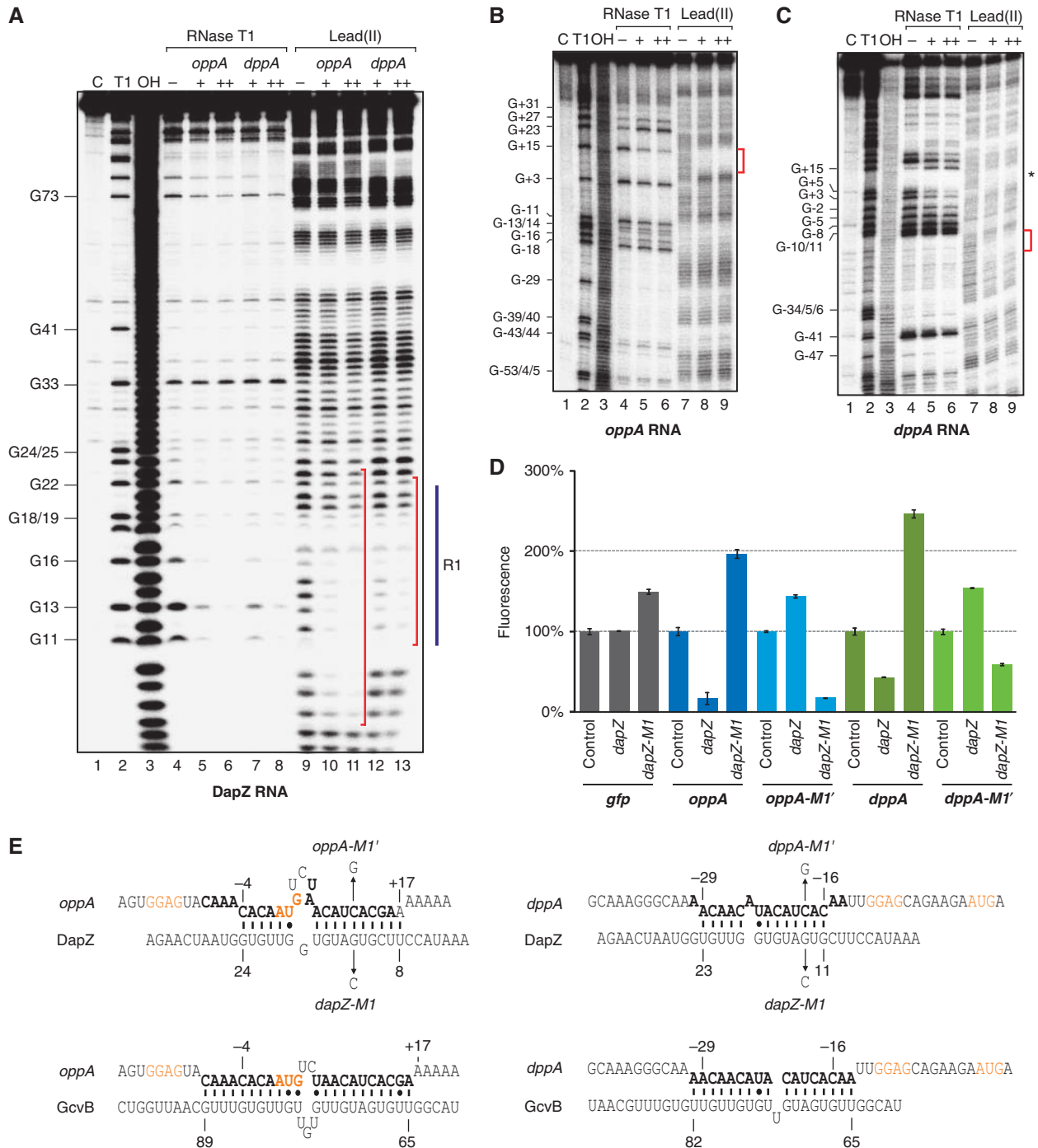


Figure 6 DapZ targets C/A-rich sites in *oppA* and *dppA*. Identification of duplex formation sites by *in vitro* secondary structure probing using 5' end-labelled DapZ sRNA (A), *oppA* (B) and *dppA* (C) mRNA leaders. Radio-labelled RNA (~5 nM) was subjected to RNase T1 or lead (II) cleavage in absence or presence of unlabelled 100 nM (+), 500 nM (++) and 20 nM (+), 100 nM (++) *dppA* in (A), 100 nM (+), 500 nM (++) DapZ in (B), and 20 nM (+), 100 nM (++) DapZ in (C). C: control RNA, T1: RNase T1 ladder, OH: alkaline ladder. The G residues are labelled relative to the translation start site in *oppA* and *dppA* mRNA leaders. The regions protected by duplex formation with cold RNA are marked with red square brackets. The DapZ R1 region is indicated with a blue bar. (*) denotes a structure rearrangement in *dppA*, which we tentatively exclude as another targeting site, because the GAGUAUUUCCUU nucleotides (+3 to +14 of *dppA*) in question have no obvious complementarity with DapZ. Thus, there might be further structural rearrangement of the mRNA upon DapZ binding, which would also explain the difference in migration of *dppA* leader RNA in native gels (Supplementary Figure S8). (D) Validation of the base-pair interactions using translational *oppA::gfp* and *dppA::gfp* reporter gene fusions by compensatory base-pair exchange *in vivo*. *Salmonella* strains containing both a *gfp* reporter plasmid and a vector overexpressing DapZ were grown overnight in LB and analysed by flow cytometry. Overexpression of a ~50 nt nonsense RNA was used as control (pJV300). (E) RNA duplexes formed between DapZ sRNA and the *dppA* or *oppA* leaders. Nucleotides in bold in *oppA* and *dppA* were previously shown to be involved in binding to GcvB sRNA (Sharma *et al*, 2007); the GcvB-*dppA* and GcvB-*oppA* interactions are shown for comparison below. Point mutations introduced for compensatory base-pair exchange experiments are indicated. The ribosome binding site and the start codon are marked in orange.

used to rule out physiological activity. For example, although SgrS RNA constitutes <0.1% of all reads in early stationary phase, it strongly regulates the mRNA of the secreted virulence factor SopD under this condition (Papenfert *et al*, 2012).

Our investigation into the narrow expression timing of DapZ uncovered a recently evolved transcriptional control by HilD (Figure 4). HilD is the master transcriptional activator of the SPI-1 invasion genes, and additional targets outside SPI-1 are typically among those ~25% of *Salmonella* genes that were horizontally acquired since *Salmonella* and *E. coli* diverged from a common ancestor (Porwollik and McClelland, 2003; Ellermeier and Slauch, 2004). Importantly, HilD activation of the *dapZ* promoter within *dapB* provides *Salmonella* with a selective repressor of major ABC transporter synthesis under invasion conditions, as compared to recruiting the GcvB sRNA whose activity would impact amino-acid uptake and synthesis in a much broader fashion. In more general terms, HilD and DapZ constitute a novel paradigm of how the core and accessory parts of the *Salmonella* genome are intermeshed by Hfq-dependent regulation at the RNA level. Contrasting previously reported sRNA-mediated mRNA control across the conserved and virulence regions (Pfeiffer *et al*, 2007; Papenfert *et al*, 2012), this study reveals novel cross-wiring wherein a horizontally acquired transcription factor (HilD) employs a conserved Hfq-dependent sRNA (DapZ) to regulate *Salmonella* core gene (*dpp* and *opp*) expression. Considering that *Salmonella* expresses >140 sRNAs (Pfeiffer *et al*, 2007; Padalon-Brauch *et al*, 2008; Sittka *et al*, 2008; Kröger *et al*, 2012), additional Hfq-dependent regulation connecting the two parts of the *Salmonella* genome will surely be discovered.

We present the most comprehensive profiling of Hfq-associated RNAs for any organism to date. This atlas of Hfq targets must now be expanded by the integration of more growth and stress conditions, and genome-wide maps of transcription start and RNA polymerase binding sites (Cho *et al*, 2009; Mooney *et al*, 2009; Kröger *et al*, 2012), to unravel the full scope of the Hfq network.

3' UTR-derived sRNAs and dual output from mRNA loci

Most mRNAs and sRNAs are encoded by loci with single output function, but exceptions from simplicity have been known. Some small genes give rise to a dual-function mRNA which both serves as the template for synthesis of a short peptide and acts as an antisense regulator on other mRNAs *in trans* (Wadler and Vanderpool, 2007; Sonnleitner *et al*, 2008; Gimpel *et al*, 2010; Romby and Charpentier, 2010). In addition, earlier studies in *E. coli* detected abundant sRNAs from UTRs (Tjaden *et al*, 2002; Vogel *et al*, 2003; Zhang *et al*, 2003; Kawano *et al*, 2005). These observations were conceptualized in a model of 'parallel transcriptional output' wherein some protein-coding genes may produce both an mRNA template for translation and a regulatory RNA, by dual use of either the promoter or terminator (Vogel *et al*, 2003). While this model recently received support by the discovery of 5' UTR (riboswitch)-derived sRNAs in *Listeria monocytogenes* that regulate a virulence factor *in trans* (Loh *et al*, 2009), DapZ represents the first example of dual output via the 3' UTR of a widely conserved gene.

3' UTR-derived sRNAs may be generated by two major biogenesis pathways (Figure 7). The first is exemplified by DapZ, which is transcribed from an mRNA-internal promoter

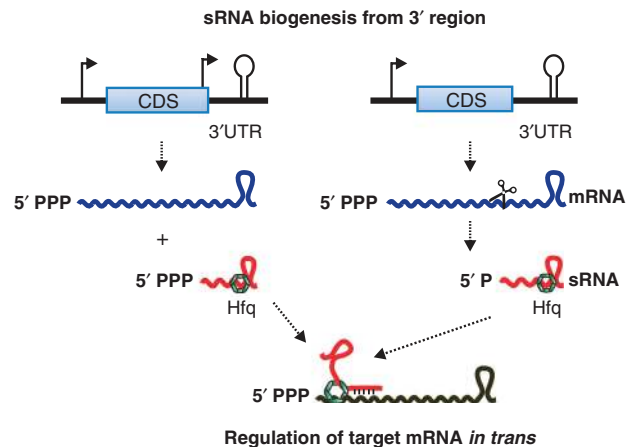


Figure 7 Biogenesis of sRNAs from bacterial UTRs. A 3' UTR-derived small RNA can be either transcribed from its own promoter in the upstream coding sequence, or generated by internal processing of the associated mRNA. The common denominator is the shared use of the ρ -independent terminator of the mRNA. Hfq plays a seminal role in either pathway such that it facilitates the base pairing of the 3' UTR-derived sRNA with *trans*-encoded target mRNA(s), but it may also participate in recruiting a nuclease (such as RNase E) to the 3' end of the mRNA in the case of processing.

and thus expressed and functions independently of the host mRNA gene despite a partial overlap in sequence. In the other pathway, the sRNA is generated by mRNA processing in the 3' region. Inspection of 5' end status in available dRNA-seq data sets wherein primary and processed transcripts can be differentiated (Sharma *et al*, 2010; Kröger *et al*, 2012) supports a notion that the 3' derived sRNAs shown in Figure 2 represent examples of both pathways: Aside from DapZ, the STnc860 sRNA (RyeF) is a primary transcript, with its promoter being located ~260 bp upstream of the *cutC* stop codon (Supplementary Figure S9). RyeF was originally discovered by Hfq coIP in *E. coli* (Zhang *et al*, 2003) but it was previously not annotated in *Salmonella* due to poorly conserved flanking genes and lack of expression in early stationary phase (Figure 2; Sittka *et al*, 2008, 2009). Overall, these sRNAs with mRNA-internal promoters bolster recent findings that transcription start sites within the coding sequences of *E. coli* and *Salmonella* may be common (Kawano *et al*, 2005; Mendoza-Vargas *et al*, 2009; Kröger *et al*, 2012).

Processing in the mRNA 3' region likely underlies the generation of STnc840, STnc870 and STnc2090, because these sRNAs seem to carry a 5' monophosphate end (according to dRNA-seq data by Kröger *et al*, 2012) and lack obvious motifs for RNAP binding in the respective upstream DNA regions. The processing must come with a cost given that these cleavages collectively occur in the coding sequence and so render the respective mRNA inactive for further rounds of translation. The RybD sRNA, which was originally detected at the 3' end of the conserved *sucABCD* operon in *E. coli* (Zhang *et al*, 2003), also belongs to this category. *Salmonella* RybD is prominent in the early stationary phase (Figure 1B) and likely generated by mRNA cleavage 12 nucleotides upstream of the *sucD* stop codon (Supplementary Figure S10). A promising candidate for the responsible nuclease is RNase E, the major endoribonuclease in Gram-negative bacteria (Belasco, 2010;

Bouvier and Carpousis, 2011) which can cleave in the 3' regions of mRNAs by either processive action from the 5' end (Mackie, 1998) or direct internal entry (Kime *et al*, 2010). Of note, RNase E was proposed to process MicX sRNA from a long 3' UTR in *Vibrio cholerae* (Davis and Waldor, 2007). RNase E and Hfq can form a complex (Morita *et al*, 2005; Ikeda *et al*, 2011), which could then be guided to the 3' end of mRNAs by the recently discovered propensity of Hfq to bind to ρ -independent terminators (Otaka *et al*, 2011; Sauer and Weichenrieder, 2011; Ishikawa *et al*, 2012). While experiments are in progress to map Hfq sites and RNase E-dependent cleavage in the new 3' derived sRNAs, we note that many Hfq-associated mRNAs show enrichment in the putative terminator region. Of the predicted 770 ρ -independent terminators of *Salmonella* mRNA genes, 291 were significantly enriched by coIP with Hfq (Supplementary Table S4), which illustrates a vast array of candidate regions for new 3' UTR-derived sRNAs.

Functional classification of sRNA by seed

An exciting finding of this paper is the discovery of a GcvB-like seed region in DapZ, with potential ramifications for sRNA ontology. Except for *E. coli* and *Shigella* where the R1 region is mutated, all predicted DapZ RNAs contain this ~15 nt long G/U-rich seed which, according to our probing experiments, targets the same C/A-rich sites in *dppA* and *oppA* as does the longer R1 domain of GcvB (Figure 6; (Sharma *et al*, 2007). Inferring from the proven inhibition of ribosome loading by GcvB at these sites (Sharma *et al*, 2007), we posit that DapZ represses the *dppA* and *oppA* mRNAs primarily at the level of translation initiation.

The concept of bacterial seed domains is recent (Storz *et al*, 2011; Vogel and Luisi, 2011), and arose through observations that the nucleotides involved in binding to target mRNAs are usually conserved (Vanderpool and Gottesman, 2004; Udekwu *et al*, 2005; Sharma *et al*, 2007) and that seed regions as small as 13 nucleotides retain their function upon transplantation to unrelated sRNAs (Bouvier *et al*, 2008b; Pfeiffer *et al*, 2009; Papenfort *et al*, 2010). The R1 domains of DapZ and GcvB strongly support the concept of independent seed regions in sRNAs, and may even constitute a case of convergent evolution. The coinciding binding regions of DapZ and GcvB in *dppA* and *oppA* may indicate that these mRNAs are constrained with respect to where sRNAs can operate effectively, e.g., in terms of target site accessibility and the presence of an Hfq site (Busch *et al*, 2008; Link *et al*, 2009; Peer and Margalit, 2011; Beisel *et al*, 2012). This observation adds to recent reports of clustered sRNA sites in the *rpoS* (Mandin and Gottesman, 2010), *csqD* (Holmqvist *et al*, 2010; Jorgensen *et al*, 2012; Mika *et al*, 2012; Thomason *et al*, 2012), *ompD* (Balbontin *et al*, 2010; Gogol *et al*, 2011; Fröhlich *et al*, 2012) and *sdhC* mRNAs (Desnoyers and Masse, 2012).

We presently tend to classify Hfq-associated sRNAs according to their cognate transcription factors (Corcoran *et al*, 2011). As more sRNAs are discovered, one may want to consider an alternative nomenclature that is based on the seed. Accordingly, DapZ would be a GcvB-like sRNA, that is similar to referring to HilD as an AraC-like transcription factor because it carries the same DNA-binding domain as the ancestral AraC protein (Schechter *et al*, 1999). Note, however, that while 'domain' in a transcription factor

denotes a structurally recognizable, self-folding entity of a protein that confers a specific function for DNA binding, the seed region is defined primarily by the specific mRNA interactions it bestows on an sRNA. A nomenclature that follows the seed would seem particularly appropriate for sRNAs like DapZ whose *Salmonella*-specific control by HilD rejects the cognate transcription factor as a useful classifier.

Materials and methods

Bacterial strains, media and growth conditions

Salmonella enterica serovar Typhimurium strain SL1344 (JVS-1574) was used as wild-type strain and for mutant construction. The complete list of bacterial strains used in this study is provided in Supplementary Table S6. Bacteria were grown in Luria Bertani (LB) medium at 37°C at 220 r.p.m. When appropriate, 100 μ g/ml ampicillin, 50 μ g/ml kanamycin or 20 μ g/ml chloramphenicol (final concentrations) were added to the medium, or used in agar plates.

Strain construction

Chromosomal mutagenesis of *Salmonella* SL1344 was performed as previously described (Datsenko and Wanner, 2000). To construct the *dapZ* deletion strain (JVS-9207), 3' UTR of *dapB* was replaced by the 'scar' sequence while the ρ -independent terminator was kept intact. Wild-type *Salmonella* containing pKD46 was electroporated with 300–500 ng DNA amplified from pKD4 with oligonucleotides JVO-7698/-5641 (see Supplementary Table S8 for sequences of deoxyoligonucleotides). The kanamycin resistance cassette was eliminated using the FLP helper plasmid pCP20 at 42°C (Datsenko and Wanner, 2000). The same strategy was applied constructing the Δ rtsAB strain using primer pair JVO-5604/-5605 (verified by PCR with JVO-5606/-5607).

For construction of translational *oppA*-17aa-*lacZ* (JVS-8992) and *dppA*-10aa-*lacZ* (JVS-8996) fusions in the chromosome, the 3' part of the *oppA* or *dppA* coding sequences were first replaced by kanamycin resistance cassette amplified with oligonucleotides JVO-7322/-7323, or JVO-7324/-7325 from pKD13. The resulting mutants were 'healed' by pCP20, and then transformed with pCE40 to generate translational *lacZ* fusions (Ellermeier *et al*, 2002). The insertion of *lacZ* gene was verified by colony PCR using oligonucleotides pMC874-lac and JVO-0421(*oppA*) or JVO-0423(*dppA*). All mutations were transduced into fresh wild-type or desired *Salmonella* background using phage P22 (Sternberg and Maurer, 1991).

Plasmid construction

A complete list of plasmids used in this study can be found in Supplementary Table S7. In order to make the *pdapZ* (pYC20) construct, *dapZ* and ~800 bp upstream sequence (until the start codon of *dapB* CDS) was amplified using oligonucleotides JVO-5373/-5374, digested with *Xba*I and *Xho*I and ligated plasmid pZE12-luc digested with the same enzymes. To construct an L-arabinose inducible *dapZ* construct (pBAD-*dapZ*, pYC39), the *dapZ* gene amplified by oligonucleotides JVO-5646/-5374 was inserted into pBAD-His-myc (Invitrogen) backbone which was amplified by oligonucleotides JVO-0900/-0901. To generate the constitutive *dapZ* overexpression plasmid (pYC40-2) driven by the P_{LlacO} promoter, the very same insert was cloned into pZE12-luc backbone, which was prepared by PCR amplification with oligonucleotides PlacB and PlacD with Phusion DNA polymerase (Finnzymes, Finland).

The 4 bp (GATG) in *dapZ* R1 region was deleted from pYC39 by overlapping PCR with oligonucleotides JVO-8992/-8993. After *Dpn*I digest, the PCR product was transformed into *E. coli* TOP10 for generating plasmid P_{BAD-dapZ- Δ R1} (pYC108). Likewise, the point mutation in *dapZ* was introduced in pYC40-2 by overlapping PCR with JVO-7197/-7198 resulting in plasmid *dapZ*-M1 (pYC73-2); the compensatory point mutation in *oppA* was introduced in pJL19-1 by overlapping PCR with JVO-7199/-7200 resulting in plasmid *oppA*-M1' (pYC74). The compensatory point mutation in *dppA* was introduced in pJL18-1 with JVO-7201/-7202 resulting in plasmid *dppA*-M1' (pYC75-3).

Transcriptional P_{*dapZ*}-*gfp* fusion plasmids were constructed by cloning DNA sequences containing putative promoter elements into

plasmid pAS0046 via *AatII/NheI* sites as previously described (Pfeiffer *et al*, 2007). DNA sequences containing putative *dapZ* promoters in enterobacteria species (~110 to +5 bp relative to the +1 site of *Salmonella dapZ*) were amplified from genomic DNA of wild-type *Salmonella* with JVO-7635/-7636, *E. coli* K-12 with JVO-7637/-7638, *Yersinia pestis* KUMA with JVO-7639/-7640, *Klebsiella pneumoniae* MGH78578 with JVO-7641/-7642 and *Citrobacter rodentium* with JVO-7643/-7644.

coIP of Hfq and deep sequencing analysis

Wild-type *Salmonella* SL1344 and the 3 × FLAG-tagged Hfq strain (*hfq*^{FLAG}, JVS-1338) were grown in LB overnight for about 16 h (220 r.p.m., 37°C). Cells equivalent to an OD of 50 were collected for coIP, and another 5 ml overnight culture was washed by PBS twice, diluted 1:200 to 1 l fresh LB and grown for 9 h after reaching OD₆₀₀ of 2. At several time points during growth (Figure 1A), a 50 OD culture of wild-type and *hfq*^{FLAG} strain were collected by centrifugation (4000 r.p.m., 4°C) and subjected to coIP according to the protocol previously described by Pfeiffer *et al* (2007) and Sittka *et al* (2008), with slight modifications. Briefly, bacteria were resuspended in 0.8 ml of lysis buffer (20 mM Tris pH 8, 150 mM KCl, 1 mM MgCl₂, 1 mM DTT), and disrupted with 0.8 ml glass beads (Roth, 0.1 mm diameter) by rigorous vortexing (30 s burst followed by 30 s chill on ice) for 5 min. The cleared lysate was incubated with 35 µl anti-FLAG antibody (Sigma; #F1804) at 4°C for 30 min and incubated with 75 µl of Protein A sepharose (Sigma; P-6649-5ML) for another 30 min. After five washes in lysis buffer, the sepharose was subjected to RNA and protein preparation by Phenol:Chloroform:Isopropanol extraction. After DNase I digestion, the RNA was used to construct cDNA libraries by Vertis Biotechnologie AG (Munich, Germany), and sequenced on a Solexa GAIIX machine. Raw cDNA reads were quality trimmed (cutoff Phread score of 20) and poly-A clipped. Solexa reads (≥20 nt) were mapped to the *Salmonella* SL1344 genome (<http://www.sanger.ac.uk/Projects/Salmonella/>) using *segemehl* with a minimal accuracy of 95% (Hoffmann *et al*, 2009). The per nucleotide coverage was calculated, normalized by the number of total number of mapped of reads for each library and visualized in the Integrated Genome Browser (Nicol *et al*, 2009). Gene-wise quantity analysis was performed by counting the number of reads which overlap at least for 10 nt with genes annotations in the set recently generated by Kröger *et al* (2012). The sequencing data have been deposited in the GEO database (accession no. GSE38884). To calculate the enrichment of RNAs, the number of reads from Hfq coIP library was divided by that from control coIP library, followed by normalization to the total number of mapped reads.

RNA isolation and northern hybridization

RNA isolation and northern hybridization experiments were performed as previously described (Pfeiffer *et al*, 2007; Sittka *et al*, 2008). Briefly, samples were collected with addition of 0.2 vol/vol of STOP solution (95% ethanol, 5% phenol) before the RNA was prepared with Trizol reagent (Invitrogen). 5–10 µg total RNA was denatured for 5 min at 95°C in RNA loading buffer (95% formamide, 0.1% xylene cyanole, 0.1% bromophenol blue and 10 mM EDTA), separated on 7 M urea/ 6% polyacrylamide gels, and transferred onto Hybond-XL membranes (GE Healthcare) by electroblotting (50 V) for 1 h at 4°C. Oligos were 5' end-labelled with γ -³²P by PNK (Fermentas). All oligonucleotides used for detection of *DapZ*, newly identified 3' UTR sRNAs (Figure 2) and 5S rRNA are listed in Supplementary Table S8. The 5' end-labelled oligos were hybridized to membranes overnight at 42°C, before washing with 5 × SSC/0.1% SDS, 1 × SSC/0.1% SDS and 0.5 × SSC/0.1% SDS for 15 min each. Signals were visualized on a phosphorimager (Typhoon FLA 7000, GE Healthcare) and quantified using the AIDA software (Raytest).

5' RACE

Primary transcription start sites were determined with a modified 5' RACE protocol (Bensing *et al*, 1996) as described in detail by Pfeiffer *et al* (2007). Briefly, 12 µg total RNA was split in two parts while only one was treated with Tobacco Acid Pyrophosphatase (TAP; Epicentre) at 37°C for 30 min. Subsequently, RNA was ligated to a 5' RNA oligonucleotide adaptor using T4 RNA ligase at 17°C overnight. Following purification, half of the ligated RNA was

used for reverse transcription with a random hexamer oligonucleotide primer mix using Superscript III RT (200 U final; Invitrogen) with the following programme: 25°C for 5 min, 50°C for 60 min, 70°C for 15 min. Oligonucleotides JVO-4661 and JVO-0367 (antisense to the RNA linker) were used to amplify the 5' end *DapZ* by PCR with Taq polymerase (New England Biolabs), following 35 cycles of: 95°C for 20 s, 56°C for 20 s, 72°C for 20 s. The PCR products were separated on a 4% agarose gel and the sequence of the TAP-specific band was determined by Sanger sequencing.

Microarray analysis

Microarray analysis of pBAD-induced sRNAs expression has been described previously (Papenfert *et al*, 2006, 2009, 2012). In brief, wild-type *Salmonella* was transformed with plasmid pKP8-35 (pBAD control), pYC39 (pBAD-*DapZ*) or pYC108 (pBAD-*DapZ*- Δ R1), and grown in LB until OD₆₀₀ of 1.5. The expression of sRNA was induced for 10 min with 0.2% L-arabinose, and total RNA was prepared for microarray analysis with Trizol reagent as described above. Differential expression was considered when genes displayed ≥2-fold changes in all replicates and were statistically significantly different (Student's *t*-test). Statistical analysis, data visualization and data mining were analysed using GeneSpring 7.3 (Agilent). The microarray data have been deposited in the GEO database (accession no. GSE38523).

Western blot and antibodies

Bacteria culture were collected by centrifugation for 2 min at 16 100 g at 4°C, and pellets were resuspended in 1 × protein loading buffer to a final concentration of 0.01 OD/µl. After incubation for 5 min at 95°C, 0.05 OD equivalents of samples were separated on 12% SDS-PAGE. GFP fusion and GroEL proteins were detected as described in Urban and Vogel (2007). HilD was detected with a polyclonal anti-HilD antibody.

β-galactosidase and GFP reporter assays

β-Galactosidase activity was determined with *ortho*-Nitrophenyl-β-galactoside (ONPG) as substrate and the CHCl₃-SDS permeabilization procedure (Miller, 1972). To assay GFP reporters, bacteria were grown in LB in presence of ampicillin and chloramphenicol overnight. Bacterial cells corresponding to 1 OD were pelleted and fixed with 4% paraformaldehyde. The GFP fluorescence intensity was quantified by flow cytometry with FACS Calibur (BD Bioscience). All experiments were performed in triplicates.

In-vitro structure mapping and foot-printing

DNA templates carrying a T7 promoter sequence for *in-vitro* transcription were generated by PCR. Primers and sequences of the T7 transcripts have been deposited in Supplementary Table S7. RNA was *in vitro* transcribed and quality checked as described (Sharma *et al*, 2007; Sittka *et al*, 2007). The protocol for 5' end labelling of RNA has been described previously (Papenfert *et al*, 2006).

Secondary structure probing and mapping of RNA duplexes were conducted with ~0.1 pmol 5' end-labelled RNA in 10 µl reactions. RNA was denatured for 1 min at 95°C and chilled on ice for 5 min, upon which 1 µg of yeast tRNA and 10 × structure buffer (0.1 M Tris at pH 7.0, 1 M KCl, 0.1 M MgCl₂; Ambion), and unlabelled mRNA/sRNA were added to anneal at 37°C for 10 min. Thereafter, 2 µl of RNase T1 (0.05 U/µl; Ambion, #AM2283) or 2 µl of a fresh solution of lead(II) acetate (25 mM; Fluka #15319) was added and incubated for additional 3 or 1 min at 37°C, respectively. Reactions were stopped by adding 12 µl cold loading buffer II (95% Formamide; 18 mM EDTA; 0.025% SDS, Xylene Cyanol and Bromophenol Blue; Ambion). RNase T1 ladders were obtained by incubating labelled RNA (~0.2 pmol) in 1 × sequencing buffer (Ambion) for 1 min at 95°C. Subsequently, 1 µl of RNase T1 (0.1 U/µl) was added, and incubation was continued for 5 min at 37°C. The OH ladder was generated by incubation of 0.2 pmol of labelled RNA for 5 min in alkaline hydrolysis buffer (Ambion) at 95°C. Samples were denatured for 3 min at 95°C prior to separation on 6% polyacrylamide/7 M urea sequencing gels in 1 × TBE. Gels were dried and analysed using PhosphorImager FLA-7000 and AIDA software.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank José L Puente for providing a *Salmonella* Δ hflD strain and Akiko Takaya for HflD antibody; Konrad Förstner for sequence analysis; Hans Mollenkopf for help with microarray experiments; vertis Biotechnologie AG (Munich, Germany) for reliable cDNA library construction; Carsten Kröger and Jay Hinton for sharing unpublished results; Stan Gorski for help with the manuscript. This work was supported by DFG Priority Program

SPP1258 Sensory and Regulatory RNAs in Prokaryotes (DFG grants Vo875/3-2 and DFG Vo875/4-2 and BMBF (German Ministry of Education and Research) grants *Next-generation transcriptomics of bacterial infections* and *RNomics in Infectious Diseases*; 01GS0806). YC was recipient of a scholarship from International Max-Planck Research School (IMPRS-IDI).

Author contributions: YC and JV designed the study; YC performed the experiments; YC, KP and CMS analysed the data; RR performed deep sequencing; YC, KP and JV wrote the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Andrade JM, Pobre V, Matos AM, Arraiano CM (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18**: 844–855
- Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, Margalit H, Altuvia S (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr Biol* **11**: 941–950
- Backofen R, Hess WR (2010) Computational prediction of sRNAs and their targets in bacteria. *RNA Biol* **7**: 33–42
- Balbotin R, Fiorini F, Figueroa-Bossi N, Casadesus J, Bossi L (2010) Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Mol Microbiol* **78**: 380–394
- Beisel CL, Storz G (2011) The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol Cell* **41**: 286–297
- Beisel CL, Updegrove TB, Janson BJ, Storz G (2012) Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J* **31**: 1961–1974
- Belasco JG (2010) All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat Rev Mol Cell Biol* **11**: 467–478
- Bensing BA, Meyer BJ, Dunny GM (1996) Sensitive detection of bacterial transcription initiation sites and differentiation from RNA processing sites in the pheromone-induced plasmid transfer system of *Enterococcus faecalis*. *Proc Natl Acad Sci USA* **93**: 7794–7799
- Bouvier J, Stragier P, Morales V, Remy E, Gutierrez C (2008a) Lysine represses transcription of the *Escherichia coli* *dapB* gene by preventing its activation by the ArgP activator. *J Bacteriol* **190**: 5224–5229
- Bouvier M, Carpousis AJ (2011) A tale of two mRNA degradation pathways mediated by RNase E. *Mol Microbiol* **82**: 1305–1310
- Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J (2008b) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* **32**: 827–837
- Boysen A, Moller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M (2010) Translational regulation of gene expression by an anaerobically induced small non-coding RNA in *Escherichia coli*. *J Biol Chem* **285**: 10690–10702
- Busch A, Richter AS, Backofen R (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**: 2849–2856
- Chen S, Lesnik EA, Hall TA, Sampath R, Griffey RH, Ecker DJ, Blyn LB (2002) A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Biosystems* **65**: 157–177
- Cho BK, Zengler K, Qiu Y, Park YS, Knight EM, Barrett CL, Gao Y, Palsson BO (2009) The transcription unit architecture of the *Escherichia coli* genome. *Nat Biotechnol* **27**: 1043–1049
- Corcoran CP, Papenfort K, Vogel J (2011) Hfq-associated regulatory small RNAs. In *Regulatory RNAs in Prokaryotes*, Marchfelder A, Hess WR (eds), pp 15–50. Springer
- Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JC, Vogel J (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol* **84**: 428–445
- Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**: 6640–6645
- Davis BM, Waldor MK (2007) RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol Microbiol* **65**: 373–385
- Deana A, Celesnik H, Belasco JG (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**: 355–358
- Desnoyers G, Masse E (2012) Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes Dev* **26**: 726–739
- Durand S, Storz G (2010) Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol Microbiol* **75**: 1215–1231
- Ellermeier CD, Janakiraman A, Schlauch JM (2002) Construction of targeted single copy lac fusions using lambda Red and FLP-mediated site-specific recombination in bacteria. *Gene* **290**: 153–161
- Ellermeier CD, Schlauch JM (2004) RtsA coordinately regulates DsbA and the *Salmonella* pathogenicity island 1 type III secretion system. *J Bacteriol* **186**: 68–79
- Ellermeier JR, Schlauch JM (2007) Adaptation to the host environment: regulation of the SPI1 type III secretion system in *Salmonella enterica* serovar Typhimurium. *Curr Opin Microbiol* **10**: 24–29
- Fender A, Elf J, Hampel K, Zimmermann B, Wagner EG (2010) RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev* **24**: 2621–2626
- Figueroa-Bossi N, Valentini M, Malleret L, Fiorini F, Bossi L (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev* **23**: 2004–2015
- Fröhlich KS, Papenfort K, Berger AA, Vogel J (2012) A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. *Nucleic Acids Res* **40**: 3623–3640
- Fröhlich KS, Vogel J (2009) Activation of gene expression by small RNA. *Curr Opin Microbiol* **12**: 674–682
- Gimpel M, Heidrich N, Mader U, Krugel H, Brantl S (2010) A dual-function sRNA from *B. subtilis*: SR1 acts as a peptide encoding mRNA on the gapA operon. *Mol Microbiol* **76**: 990–1009
- Gogol EB, Rhodius VA, Papenfort K, Vogel J, Gross CA (2011) Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc Natl Acad Sci USA* **108**: 12875–12880
- Guillier M, Gottesman S (2008) The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res* **36**: 6781–6794
- Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, Vogel J, Stadler PF, Hacker Muller J (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol* **5**: e1000502
- Holmqvist E, Reimegard J, Sterk M, Grantcharova N, Romling U, Wagner EG (2010) Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J* **29**: 1840–1850
- Holmqvist E, Unoson C, Reimegard J, Wagner EG (2012) A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol Microbiol* **84**: 414–427
- Ikeda Y, Yagi M, Morita T, Aiba H (2011) Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli*. *Mol Microbiol* **79**: 419–432
- Ishikawa H, Otaka H, Maki K, Morita T, Aiba H (2012) The functional Hfq-binding module of bacterial sRNAs consists of a

- double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. *RNA* **18**: 1062–1074
- Jorgensen MG, Nielsen JS, Boysen A, Franch T, Moller-Jensen J, Valentin-Hansen P (2012) Small regulatory RNAs control the multi-cellular adhesive lifestyle of *Escherichia coli*. *Mol Microbiol* **84**: 36–50
- Kawano M, Reynolds AA, Miranda-Rios J, Storz G (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res* **33**: 1040–1050
- Kime L, Jourdan SS, Stead JA, Hidalgo-Sastre A, McDowall KJ (2010) Rapid cleavage of RNA by RNase E in the absence of 5' monophosphate stimulation. *Mol Microbiol* **76**: 590–604
- Kingsford CL, Ayanbule K, Salzberg SL (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* **8**: R22
- Kröger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, Colgan A, Leekitcharoenphon P, Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ, Thomson NR, Vogel J *et al* (2012) The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci USA* **109**: E1277–E1286
- Lease RA, Woodson SA (2004) Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol* **344**: 1211–1223
- Lesnik EA, Sampath R, Levene HB, Henderson TJ, McNeil JA, Ecker DJ (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res* **29**: 3583–3594
- Link TM, Valentin-Hansen P, Brennan RG (2009) Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci USA* **106**: 19292–19297
- Livny J, Brencic A, Lory S, Waldor MK (2006) Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. *Nucleic Acids Res* **34**: 3484–3493
- Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, Repoila F, Buchrieser C, Cossart P, Johansson J (2009) A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* **139**: 770–779
- Mackie GA (1998) Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**: 720–723
- Maki K, Morita T, Otaka H, Aiba H (2010) A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. *Mol Microbiol* **76**: 782–792
- Mandin P, Gottesman S (2010) Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* **29**: 3094–3107
- Masse E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci USA* **99**: 4620–4625
- Masse E, Vanderpool CK, Gottesman S (2005) Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* **187**: 6962–6971
- Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juarez K, Contreras-Moreira B, Huerta AM, Collado-Vides J, Morett E (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS ONE* **4**: e7526
- Mika F, Busse S, Possling A, Berkholz J, Tschowri N, Sommerfeldt N, Pruteanu M, Hengge R (2012) Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in *Escherichia coli*. *Mol Microbiol* **84**: 51–65
- Miller JH (1972) *Experiments in Molecular Genetics*. NY: Cold Spring Harbor Laboratory Press
- Moon K, Gottesman S (2009) A PhoQ/P-regulated small RNA regulates sensitivity of *Escherichia coli* to antimicrobial peptides. *Mol Microbiol* **74**: 1314–1330
- Moon K, Gottesman S (2011) Competition among Hfq-binding small RNAs in *Escherichia coli*. *Mol Microbiol* **82**: 1545–1562
- Mooney RA, Davis SE, Peters JM, Rowland JL, Ansari AZ, Landick R (2009) Regulator trafficking on bacterial transcription units *in vivo*. *Mol Cell* **33**: 97–108
- Morita T, Maki K, Aiba H (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* **19**: 2176–2186
- Nicol JW, Helt GA, Blanchard Jr. SG, Raja A, Loraine AE (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**: 2730–2731
- Otaka H, Ishikawa H, Morita T, Aiba H (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc Natl Acad Sci USA* **108**: 13059–13064
- Padalon-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, Margalit H, Altuvia S (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res* **36**: 1913–1927
- Papenfort K, Bouvier M, Mika F, Sharma CM, Vogel J (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci USA* **107**: 20435–20440
- Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, Vogel J (2006) SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Mol Microbiol* **62**: 1674–1688
- Papenfort K, Podkaminski D, Hinton JC, Vogel J (2012) The ancestral SgrS RNA discriminates horizontally acquired *Salmonella* mRNAs through a single G-U wobble pair. *Proc Natl Acad Sci USA* **109**: E757–E764
- Papenfort K, Said N, Welsink T, Lucchini S, Hinton JC, Vogel J (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* **74**: 139–158
- Peer A, Margalit H (2011) Accessibility and evolutionary conservation mark bacterial small-rna target-binding regions. *J Bacteriol* **193**: 1690–1701
- Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**: 840–846
- Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, Vogel J (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol Microbiol* **66**: 1174–1191
- Porwollik S, McClelland M (2003) Lateral gene transfer in *Salmonella*. *Microbes Infect* **5**: 977–989
- Pulvermacher SC, Stauffer LT, Stauffer GV (2008) The role of the small regulatory RNA GcvB in GcvB/mRNA posttranscriptional regulation of oppA and dppA in *Escherichia coli*. *FEMS Microbiol Lett* **281**: 42–50
- Pulvermacher SC, Stauffer LT, Stauffer GV (2009) Role of the sRNA GcvB in regulation of cycA in *Escherichia coli*. *Microbiology* **155**(Pt 1): 106–114
- Rasmussen AA, Johansen J, Nielsen JS, Overgaard M, Kallipolitis B, Valentin-Hansen P (2009) A conserved small RNA promotes silencing of the outer membrane protein YbfM. *Mol Microbiol* **72**: 566–577
- Rice JB, Vanderpool CK (2011) The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. *Nucleic Acids Res* **39**: 3806–3819
- Rivas E, Eddy SR (2001) Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics* **2**: 8
- Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**: 1902–1910
- Romy P, Charpentier E (2010) An overview of RNAs with regulatory functions in gram-positive bacteria. *Cell Mol Life Sci* **67**: 217–237
- Sauer E, Weichenrieder O (2011) Structural basis for RNA 3'-end recognition by Hfq. *Proc Natl Acad Sci USA* **108**: 13065–13070
- Schechter LM, Damrauer SM, Lee CA (1999) Two AraC/XylS family members can independently counteract the effect of repressing sequences upstream of the hilA promoter. *Mol Microbiol* **32**: 629–642
- Schechter LM, Lee CA (2001) AraC/XylS family members, HilC and HilD, directly bind and derepress the *Salmonella typhimurium* hilA promoter. *Mol Microbiol* **40**: 1289–1299
- Shao Y, Bassler BL (2012) Quorum-sensing non-coding small RNAs use unique pairing regions to differentially control mRNA targets. *Mol Microbiol* **83**: 599–611
- Sharma CM, Darfeuille F, Plantinga TH, Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich

- elements inside and upstream of ribosome-binding sites. *Genes Dev* **21**: 2804–2817
- Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hacker-muller J, Reinhardt R, Stadler PF, Vogel J (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**: 250–255
- Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf HJ, Hinton JC, Vogel J (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* **81**: 1144–1165
- Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JC, Vogel J (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* **4**: e1000163
- Sittka A, Pfeiffer V, Tedin K, Vogel J (2007) The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*. *Mol Microbiol* **63**: 193–217
- Sittka A, Sharma CM, Rolle K, 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
- Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R, Williams P, Huttenhofer A, Haas D, Blasi U (2011) The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol Microbiol* **80**: 868–885
- Sonnleitner E, Sorger-Domenigg T, Madej MJ, Findeiss S, Hacker-muller J, Huttenhofer A, Stadler PF, Blasi U, Moll I (2008) Detection of small RNAs in *Pseudomonas aeruginosa* by RNomics and structure-based bioinformatic tools. *Microbiology* **154**(Part 10): 3175–3187
- Sternberg NL, Maurer R (1991) Bacteriophage-mediated generalized transduction in *Escherichia coli* and *Salmonella typhimurium*. *Methods Enzymol* **204**: 18–43
- Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* **43**: 880–891
- Thomason MK, Fontaine F, De Lay N, Storz G (2012) A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Mol Microbiol* **84**: 17–35
- Tjaden B, Saxena RM, Stolyar S, Haynor DR, Kolker E, Rosenow C (2002) Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Res* **30**: 3732–3738
- Udekwi KI, Darfeuille F, Vogel J, Reimegard J, Holmqvist E, Wagner EGH (2005) Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev* **19**: 2355–2366
- Urban JH, Vogel J (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res* **35**: 1018–1037
- Vanderpool CK, Gottesman S (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **54**: 1076–1089
- Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jager JG, Huttenhofer A, Wagner EG (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res* **31**: 6435–6443
- Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. *Nat Rev Microbiol* **9**: 578–589
- Vogel J, Sharma CM (2005) How to find small non-coding RNAs in bacteria. *Biol Chem* **386**: 1219–1238
- Wadler CS, Vanderpool CK (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci USA* **104**: 20454–20459
- Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* **15**: 1637–1651
- Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50**: 1111–1124