RNA Biology 10:3, 397–405; March 2013; © 2013 Landes Bioscience

Identification and characterization of small RNAs in *Yersinia pestis*

Arthur Beauregard,[1] Eric A. Smith,[1] Brianna L. Petrone,[1] Navjot Singh,[1] Christopher Karch,[2] Kathleen A. McDonough[1,2,*] and Joseph T. Wade[1,2,*]

[1]Wadsworth Center; New York State Department of Health; Albany, NY USA; [2]Department of Biomedical Sciences; School of Public Health; University at Albany; Albany, NY USA

Keywords: Yersinia pestis, Y. pseudotuberculosis, plague, Hfq, small RNA, RNA-seq, Deep-RACE

Yersinia pestis, the etiologic agent of plague, is closely related to *Yersinia pseudotuberculosis* evolutionarily but has a very different mode of infection. The RNA-binding regulatory protein, Hfq, mediates regulation by small RNAs (sRNAs) and is required for virulence of both *Y. pestis* and *Y. pseudotuberculosis*. Moreover, Hfq is required for growth of *Y. pestis*, but not *Y. pseudotuberculosis*, at 37°C. Together, these observations suggest that sRNAs play important roles in the virulence and survival of *Y. pestis*, and that regulation by sRNAs may account for some of the differences between *Y. pestis* and *Y. pseudotuberculosis*. We have used a deep sequencing approach to identify 31 sRNAs in *Y. pestis*. The majority of these sRNAs are not conserved outside the Yersiniae. Expression of the sRNAs was confirmed by Northern analysis and we developed deep sequencing approaches to map 5' and 3' ends of many sRNAs simultaneously. Expression of the majority of the sRNAs we identified is dependent upon Hfq. We also observed temperature-dependent effects on the expression of many sRNAs, and differences in expression patterns between *Y. pestis* and *Y. pseudotuberculosis*. Thus, our data suggest that regulation by sRNAs plays an important role in the adaptation to both flea vector and mammalian host, and that regulation by sRNAs may contribute to the phenotypic differences between *Y. pestis* and *Y. pseudotuberculosis*.

Introduction

Yersinia pestis, the etiologic agent of plague, continues to pose a threat to human health both naturally and as a bioweapon. *Y. pestis* is closely related to *Yersinia pseudotuberculosis*. Both species are human pathogens and are believed to have diverged from each other ~1,500–20,000 y ago;[1,2] ~75% of their genes share ≥ 97% nucleotide identity.[3] Despite their genetic similarity, the diseases caused by these two organisms vary greatly. *Y. pestis* infects both mammalian and arthropod hosts and is typically transmitted to humans through the bite of an infected flea. In humans, infection by *Y. pestis* usually manifests itself as bubonic and pneumonic plague.[4] In contrast, *Y. pseudotuberculosis* is an enteropathogen that causes a gastrointestinal disease transmitted by the fecal-oral route.[5]

Small RNAs (sRNAs) serve as important components of many regulatory circuits in bacteria. sRNAs are typically non-coding RNA molecules of < 500 nt, transcribed from intergenic regions.[6] The majority of sRNAs characterized to date have been shown to downregulate gene expression at the post-transcriptional level by base-pairing with target mRNAs. These pairing interactions result in changes in transcription attenuation, translation initiation or mRNA stability.[6]

Hfq is an RNA-binding protein that is required for the stability and/or regulatory function of many sRNAs.[7] Hfq is also required for the virulence of many pathogenic bacteria,[8] including *Y. pestis*[9] and *Y. pseudotuberculosis*,[10] suggesting that sRNAs are key regulators of virulence genes in these species. Moreover, Hfq is required for efficient biofilm formation and gut blockage in the flea, important processes for transmission to a mammalian host,[11] and for the growth of some *Y. pestis* strains, but not *Y. pseudotuberculosis*, at 37°C.[12] Strikingly, Hfq amino acid sequence is 100% identical across eight sequenced strains of *Y. pestis* and four of *Y. pseudotuberculosis*, suggesting that regulation by sRNAs rather than Hfq itself contributes to the difference in viability of *Y. pestis* and *Y. pseudotuberculosis hfq* mutants at 37°C.

Most studies of sRNAs have focused on *Escherichia coli*, for which > 100 sRNAs have been identified,[13-15] and deep sequencing approaches have led to the identification of similar numbers in other bacterial species.[16-22] Although all bacterial species are expected to express numerous sRNAs, conservation of sRNAs is generally poor, so the specific sRNA pool differs widely between species. Until recently, very few sRNAs had been identified in any *Yersinia* species. A recent study utilized a deep sequencing approach to identify 150 putative sRNAs in *Y. pseudotuberculosis*.[23] The majority of the putative sRNAs identified are conserved in *Y. pestis*, but the expression and dependence on Hfq of five sRNAs (seven were tested) differs between the closely related species. The authors also showed that deletion of specific sRNAs in *Y. pseudotuberculosis* leads to attenuation of the pathogen in a mouse model of infection and that the inactivation of an sRNA in *Y. pestis* reduces virulence in a mouse model of pneumonic plague.[23]

In this work, we utilized a deep sequencing approach to identify putative sRNAs expressed in *Y. pestis*. We confirmed expression of 31 sRNAs by Northern analysis, of which only 17 match

*Correspondence to: Kathleen A. McDonough; Email: kathleen.mcdonough@wadsworth.org; Joseph T. Wade; Email: jwade@wadsworth.org
Submitted: 07/30/12; Revised: 01/10/13; Accepted: 01/13/13
http://dx.doi.org/10.4161/rna.23590

previously identified putative sRNAs.[23] We developed genome-scale 5' and 3' RACE (rapid amplification of cDNA ends) approaches to map the 5' and 3' ends of most of the sRNAs. We observed a wide variety of expression patterns that depend upon temperature and the presence of Hfq. All of the sRNAs we identified are conserved in *Y. pseudotuberculosis* (most are 100% identical), most are conserved in other *Yersinia* species, but fewer than half are conserved in *E. coli*. We observed detectable expression in *Y. pseudotuberculosis* for all but one of the sRNAs, but the temperature- and Hfq-dependent expression patterns of many sRNAs differed between *Y. pestis* and *Y. pseudotuberculosis*. Thus, our data suggest that differences in sRNA expression may contribute to the differences in *Y. pestis* and *Y. pseudotuberculosis* biology.

Identification of putative sRNAs in *Y. pestis*. To identify novel *Y. pestis* sRNAs, we purified RNA from *Y. pestis* KIM6+ grown at 37°C and constructed a cDNA library for Illumina sequencing. Following sequencing and mapping of reads to the reference genome, we identified genomic regions with contiguous sequence reads that partially or fully overlap an intergenic region (JCVI genome annotation), with at least one position having > 500 mapped sequence reads. Thus, we generated a list of 50 putative sRNAs with a high level of confidence. We excluded repetitive sequence, although we noted that many sequences mapped to repetitive sequence partially overlapping predicted transposases. Fully antisense RNAs could not be identified due to the lack of strand information in the sequencing data.

Validation and characterization of sRNAs by northern blot. To confirm the presence of the putative sRNAs, and to determine their expression profiles in *Y. pestis* and *Y. pseudotuberculosis*, we isolated RNA from both species at 28°C and 37°C in *hfq*$^+$ cells, isogenic Δ*hfq* mutants and Δ*hfq* strains complemented with a multi-copy plasmid that encodes *hfq*.[12] We then performed Northern analysis with radiolabeled oligonucleotides designed to probe each sRNA. The deep sequencing data did not provide strand information so sRNAs were first probed on the plus strand, and any that could not be detected were then probed on the minus strand. This approach confirmed that 32 of the 50 putative sRNAs are expressed at a detectable level and have a size consistent with that of an sRNA (< 500 nt). All Northern-confirmed sRNAs are listed in **Table 1** and shown in **Figure S1**. Representative examples of the northern blots are shown in **Figure 1A**. Confirmed sRNAs were assigned "Ysr" (*Yersinia* sRNA) names, in accordance with previously identified sRNAs in *Y. pseudotuberculosis*.[23] One putative sRNA is in fact a protein-coding mRNA for the gene *rmf* (see below). Of the 31 confirmed sRNAs, 14 have not been described previously, and of the remaining 17, only five have been detected by a method other than deep sequencing.[23]

We observed a remarkable variety of expression patterns with respect to temperature, dependence upon *hfq* and species. We used an unsupervised learning algorithm to group the sRNAs into seven clusters, based on their expression patterns (**Fig. 1B**). These clusters highlight expression patterns that are common to multiple sRNAs. Clusters 1 and 2 consist largely of sRNAs that are constitutively expressed in both species, regardless of temperature or the presence of *hfq* (Ysr155/RyfD, Ysr156/Ffs, Ysr161, Ysr163, Ysr177, Ysr182/6S RNA, Ysr183/SroG, Ysr146.2/187, Ysr151/RnpB, Ysr88/152, Ysr73/169, Ysr65/175 and Ysr186/CsrC). Cluster 3 consists of sRNAs that are expressed similarly in both species but whose expression is dependent upon the presence of *hfq* (Ysr145/157, Ysr159/CyaR, Ysr164 and Ysr149/181). Cluster 4 consists of sRNAs that are expressed in both species but whose expression is dependent upon the presence of *hfq* only in *Y. pestis* (Ysr151/RnpB, Ysr148/153/GlmZ, Ysr7/154/MicA and Ysr158; the protein-coding RNA, Ysr173/*rmf*, also fell in Cluster 4). Cluster 5 consists of RNAs that are expressed in both species, whose expression in *Y. pestis* is dependent upon the presence of *hfq*, and whose expression is higher at 37°C than 28°C (Ysr167, Ysr170, Ysr171 and Ysr174). We also observed sRNAs whose expression increased in the absence of *hfq* in *Y. pseudotuberculosis* but not *Y. pestis* (Ysr23/160 and Ysr165 from Cluster 6), whose expression in both species increased in the absence of *hfq* at 28°C but decreased in the absence of *hfq* at 37°C (Ysr179/CsrB from Cluster 6) or whose expression was only detectable in *Y. pestis* (Ysr172, the sole member of Cluster 7).

Interestingly, for sRNAs in Cluster 5, as well as Ysr158, Ysr173/*rmf* (Cluster 4) and Ysr165 (Cluster 6), deletion of *hfq* in *Y. pseudotuberculosis* had no substantial effect on sRNA levels whereas expression of *hfq* from a multi-copy plasmid in the Δ*hfq* strain resulted in a substantial decrease in sRNA levels relative to *hfq*$^+$ (**Fig. 1B**). This result was observed for independent biological replicates and may be due to aberrant effects of Hfq overexpression. Consistent with this, probing the same membranes with radiolabeled oligonucleotide specific to *hfq* mRNA revealed that *hfq* is grossly overexpressed in plasmid-complemented *Y. pseudotuberculosis* but not *Y. pestis* (**Fig. 1A**; **Fig. S2**). The *hfq* northern blot data also indicated that *hfq* transcript levels are substantially lower at 37°C than at 28°C in both *Y. pestis* and *Y. pseudotuberculosis*, suggesting that varying Hfq levels may contribute to temperature-dependent changes in expression of some sRNAs.

Effects of *hfq* and temperature on sRNA levels in other *Y. pestis* and *Y. pseudotuberculosis* strains. To determine whether the different effects of *hfq* and temperature on sRNA expression between *Y. pestis* and *Y. pseudotuberculosis* are species-specific rather than strain-specific, we measured expression of three sRNAs, Ysr170, Ysr172 and Ysr179/CsrB, in another *Y. pestis* strain (CO92) and three other *Y. pseudotuberculosis* strains (PTB51c, PTB57c and PTB54c; **Fig. 1C**). The effect of temperature on sRNA expression was consistent across all strains. Specifically, in both *Y. pestis* and *Y. pseudotuberculosis*, expression of Ysr170 is higher at 37°C than at 28°C, whereas expression of Ysr179 is higher at 28°C than at 37°C (expression of Ysr172 is unaffected by temperature). There are species-specific differences in expression patterns for two of the sRNAs that were observed for all strains tested: the temperature dependence of Ysr179 expression is greater in *Y. pseudotuberculosis* than in *Y. pestis* (**Fig. 1A and C**), and Ysr172 is only expressed in *Y. pestis*. In contrast, the effect of the Δ*hfq* mutation in *Y. pestis* was not completely consistent between the KIM and CO92 strains. Specifically, Ysr170 expression is less dependent upon *hfq* in *Y. pestis* CO92 than in KIM, Ysr172 expression is not dependent upon *hfq* in *Y. pestis* CO92 (fully dependent in KIM), and *hfq* suppresses Ysr179 expression at 28°C in *Y. pestis* KIM but not CO92. Thus, some of the differences in sRNA expression between

Table 1. List of validated sRNAs

sRNA name[a]	Start coordinate[b]	End coordinate[b]	Gene context[c]
Ysr151/RnpB	125535	125684	y0115 > / < / < y0117
Ysr88/152	143205	**143274**	y0129 (gltD) > / > /y0130 >
Ysr148/153/GlmZ	**421034**	**421106, 421107**	< y0376/ > /y0377 (hemY) >
Ysr7/154/MicA	**999263**	**999189**	y0887 (gshA) > / < /y0888 >
Ysr155/RyfD	1021180	**1021317**	y0913 > / > /y0914 (clpB) >
Ysr156/Ffs	**1179191**	**1179276**	< y1045/ > / < y1046 (hha)
Ysr145/157	**1229994**	**1230059, 1230060**	< y1090/ > / < y1091
Ysr158	1341830	**1341894**	y1196 (ubiF) > / > / < yt030 (tRNA-Gln)
Ysr159/CyaR	**1527324**	**1527218**	< y1378/ < / < y1379
Ysr23/160	**1593088, 1593089**	**1593141, 1593147, 1593149**	y1436 > / > / < y1437
Ysr161	**1624496**	**1624402, 1624433**	< y1465 (nanT)/ < /y1466 (cysP) >
Ysr163	1902373	1902412	< y1720/ > /y1722 >
Ysr164	**1916071, 1916098**	**1915948**	y1731 > / < /y1732 (ypeR) >
Ysr165	**1961869**	**1961985**	y1782 > / > /y1783 >
Ysr11/166/FnrS	2195768	**2195887**	y1994 (zntB) > / > / < y1995
Ysr167	2363578	2363674	< y2138/ > /y2139 (rstA) >
Ysr73/169	**2457081**	**2457002**	< y2228/ < / < y2229
Ysr170	**2550744**	**2550868, 2550869**	y2316 > / > /y2317 >
Ysr171	2815750	**2815620**	y2553 (manZ) > / < /y2554 >
Ysr172	2841229	**2841287, 2841312**	< y2579/ > / < y2580
Ysr174	**3097172**	**3097124**	< y2796 (serS)/ < / < y2797
Ysr65/175	3111993	**3111905**	< y2808/ < / < y2809 (clpA)
Ysr177	3330179	**3330055**	< y3023 (moaA)/ < /y3024 >
Ysr179/CsrB	3457464	3457752	< y3145/ < / < y3146 (syd)
Ysr45/180/GcvB	**3467030**	**3467076**	< y3154 (gcvA)/ > /y3155 >
Ysr149/181	**3501524, 3501536**	**3501452, 3501453**	< y3181 (aas)/ < / < y3182
Ysr182/6S RNA	**3632001, 3632002**	**3632114**	y3299 > / > /y3300 >
Ysr183/SroG	3904566	**3904637**	< y3520 (ribB)/ < /y3522 >
Ysr185/Spot 42	**4230754**	**4231124**	< y3808/ < /y3809 (engB) >
Ysr186/CsrC	**4232923**	4233049	< y3810/ > / < y3811 (polA)
Ysr146.2/187	**4314248**	**4314315**	< y3871/ > /y3872 (ompR) >

[a]For sRNAs identified by Koo et al., the existing name is indicated first, followed by our systematic name. Names of *E. coli* homologs are also indicated where appropriate, e.g., Ysr148/153/GlmZ was identified by Koo et al. as Ysr148, our systematic name is Ysr153, and it is homologous to *E. coli* GlmZ. Underlined names indicated sRNAs that have not been identified previously in *Yersinia* species. [b]Bold text indicates ends identified using Deep RACE techniques. Other end coordinates are estimates derived manually from the RNA-seq data. [c]Underlined genes are overlapped by the corresponding sRNA.

Y. pestis and *Y. pseudotuberculosis* are likely due to strain-specific effects. Consistent with this, our previous study revealed differences in the growth dependence on *hfq* between *Y. pestis* KIM and *Y. pestis* CO92.[12] Nevertheless, there are clear differences in sRNA expression patterns between *Y. pestis* and *Y. pseudotuberculosis* that are conserved in all strains tested.

We have previously shown that *hfq* is important for growth of *Y. pestis* but not *Y. pseudotuberculosis*, when cells are cultured at 37°C but not at 28°C.[12] Hence, any sRNAs that show expression differences between 28–37°C, between wild-type and Δ*hfq* cells or between *Y. pestis* and *Y. pseudotuberculosis*, are potentially associated with the unique biology of *Y. pestis*. As described above, we observed many temperature-dependent differences in sRNA expression. In some cases, the effect of deleting *hfq* was specific to one temperature, e.g., Ysr179/CsrB (**Fig. 1**), indicating complex interactions between temperature and *hfq* dependence. Differences in expression of sRNAs in *Y. pestis* between 28–37°C could contribute to differences in gene expression for bacteria in a flea vector and bacteria in a mammalian host. Furthermore, differences in sRNA expression patterns between *Y. pestis* and *Y. pseudotuberculosis* could contribute to the physiological differences between these species. This is particularly likely for Ysr172, which is undetectable in *Y. pseudotuberculosis* (**Fig. 1A and C**).

Mapping of sRNA ends using Deep RACE. We developed genome-scale 5' and 3' RACE approaches to precisely determine the ends of the sRNAs confirmed by Northern analysis (**Fig. 2A**

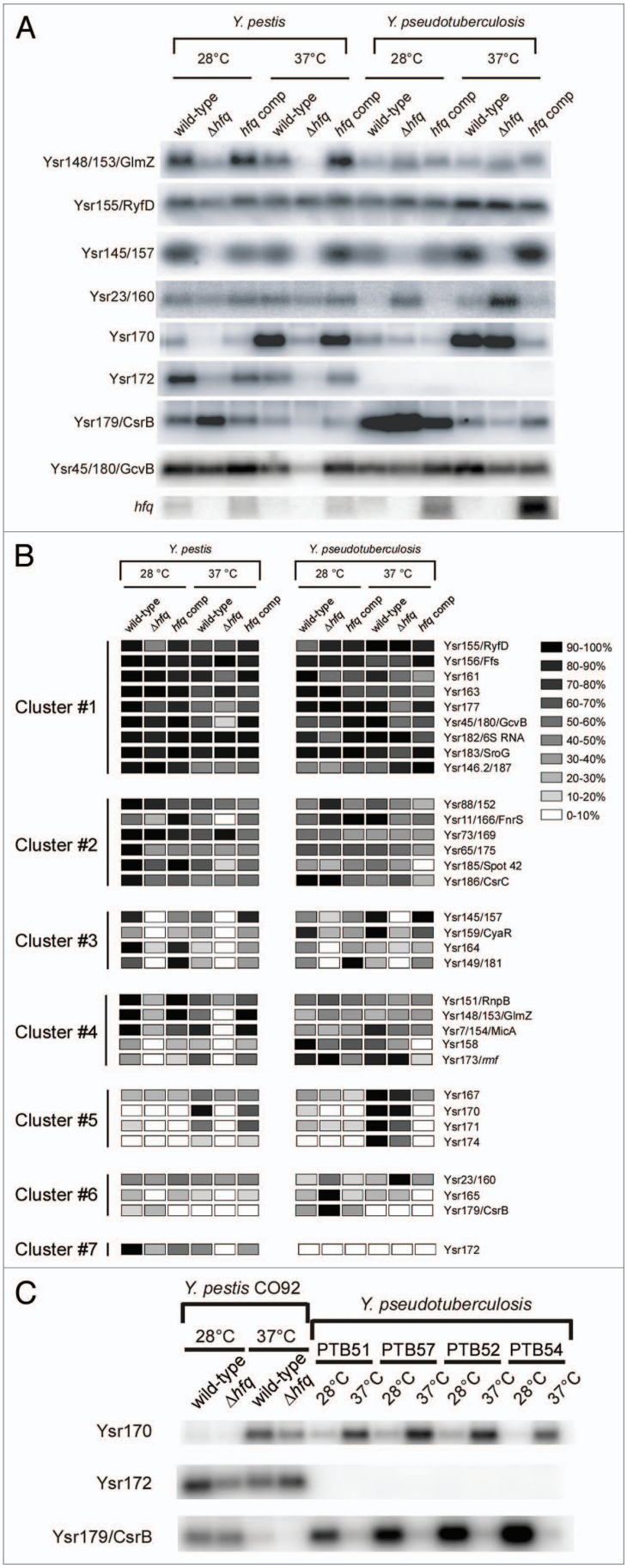

Figure 1. (**A**) Verification of Ysr expression by northern blot analysis. All northern blots are shown in **Figure S1** and duplicate northern blots for most sRNAs are shown in **Figure S4**. A northern blot for *hfq* mRNA from a corresponding set of RNA samples is also shown (note that this blot is reproduced as part of **Fig. S2**). Corresponding 5S rRNA northern blots for four of the same membranes are shown in **Figure S6**. (**B**) Clusters of sRNA based on *k*-means clustering of sRNA expression data. Rows correspond to individual RNAs, while columns correspond to conditions for which expression was measured by northern blot. Shading indicates the relative expression of sRNAs for each strain/condition. Expression numbers are indicated as a percentage of the level for condition in which the RNA level is highest. (**C**) Northern analysis of Ysr170, Ysr172 and Ysr179/CsrB in additional strains of *Y. pestis* and *Y. pseudotuberculosis*. Duplicate northern blots are shown in **Figure S5**. Note that *Y. pseudotuberculosis* PTB52c is the same strain as used in (**A**).

and B) because our deep sequencing data do not allow for such precise mapping. These methods combine conventional RACE with deep sequencing using the Ion Torrent platform (any deep sequencing platform would suffice). In addition to allowing for simultaneous analysis of many RNAs, these methods produce multiple sequence reads for each individual sRNA (e.g., 1,945 sequence reads for Ysr23/160). This allows us to identify multiple 5' ends and to accurately determine the relative abundance of each. We propose that these methods be named "Deep 5' RACE" and "Deep 3' RACE." Using these methods, we successfully mapped the 5' ends of 18 sRNAs (and *rmf* mRNA) and the 3' ends of 28 sRNAs (and *rmf* mRNA). The major 5' and 3' ends for these sRNAs are listed in **Table 1** and raw data are provided in **Tables S1 and 2**. Four representative examples are shown in **Figure 2C–F**. For the sRNAs for which we mapped both unique 5' and 3' ends, the median length is 84 nt. In most cases we detected unique ends, but some RNAs have multiple 5' ends, e.g., Ysr149/181 (**Fig. 2E**; **Table 1**). In addition, many sRNAs have multiple 3' ends clustered around a single location, e.g., Ysr148/153/GlmZ (**Fig. 2C**), Ysr149/181 (**Fig. 2E**), Ysr17/154/MicA (**Fig. 2F**; **Table 1**). Most sRNAs are located entirely within intergenic regions but some overlap the ends of adjacent genes, e.g., Ysr165 (**Fig. 2D**; **Table 1**). Our Deep 5' RACE method is very similar to a previously described method, "Deep-RACE."[24] To the best of our knowledge, no method equivalent to Deep 3' RACE has been described previously. Given the increasing availability of deep sequencing, we anticipate that these methods will become widespread for the large-scale identification of RNA 5' and 3' ends.

***Y. pestis* sRNAs fall into multiple classes based on overlap with annotated genes.** Mapping of 5' and 3' ends revealed multiple classes of sRNA. The major class is intergenic, i.e., no overlap with annotated genes. Based on our knowledge of equivalent sRNAs in other bacterial species, we anticipate that the majority of intergenic sRNAs function as regulators by base-pairing with distally encoded mRNAs. In contrast to the intergenic sRNAs, seven of the sRNAs overlap an annotated protein-coding gene. In some cases, this may be an artifact of incorrect gene annotation.

However, several of the overlapped genes have well-described functions. Two of the sRNAs, Ysr167 and Ysr171, overlap annotated genes in the antisense orientation. These sRNAs may be responsible for regulation of the overlapping gene, as has been observed previously for antisense RNAs in other species.[25] Four of the sRNAs, Ysr88/152, Ysr155/RyfD, Ysr161 and Ysr165, overlap the 5' end of an annotated gene, in the sense orientation. These sRNAs may include riboswitches which can generate short RNAs at the start of genes by promoting transcription attenuation or RNA processing. One sRNA, Ysr73/169, overlaps the 3' end of a gene in the sense orientation. This, and other sRNAs that overlap annotated genes in the sense orientation, may be processed fragments of the mRNAs that they overlap.

We determined whether any of the sRNAs might be protein-coding. Specifically, we translated sRNAs in silico and searched for open reading frames of > 30 amino acids that have significant sequence identity with proteins annotated in other species. We identified one RNA, initially named Ysr173, which encodes a homolog of *E. coli* Rmf, ribosome modulation factor.[26] We note that *rmf* is not annotated for *Y. pestis* but is annotated for *Y. pseudotuberculosis*. Other sRNAs might also be protein-coding, as has been observed for some sRNAs in *E. coli*.[27] This is particularly true for *Y. pestis*, which has a less well-annotated genome as compared with *E. coli*. Indeed, we found many differences between the gene annotations available for *Y. pestis* KIM from different databases. Strikingly, *rmf* mRNA levels in *Y. pestis* (but not *Y. pseudotuberculosis*) are dependent upon the presence of *hfq* (**Fig. 1B**; **Fig. S1**), indicating that the effects of Hfq specific to *Y. pestis* are not limited to non-coding RNAs. *rmf* abundance may be controlled by direct association of Hfq. Alternatively, *rmf* may be regulated by an sRNA in an Hfq-dependent manner.

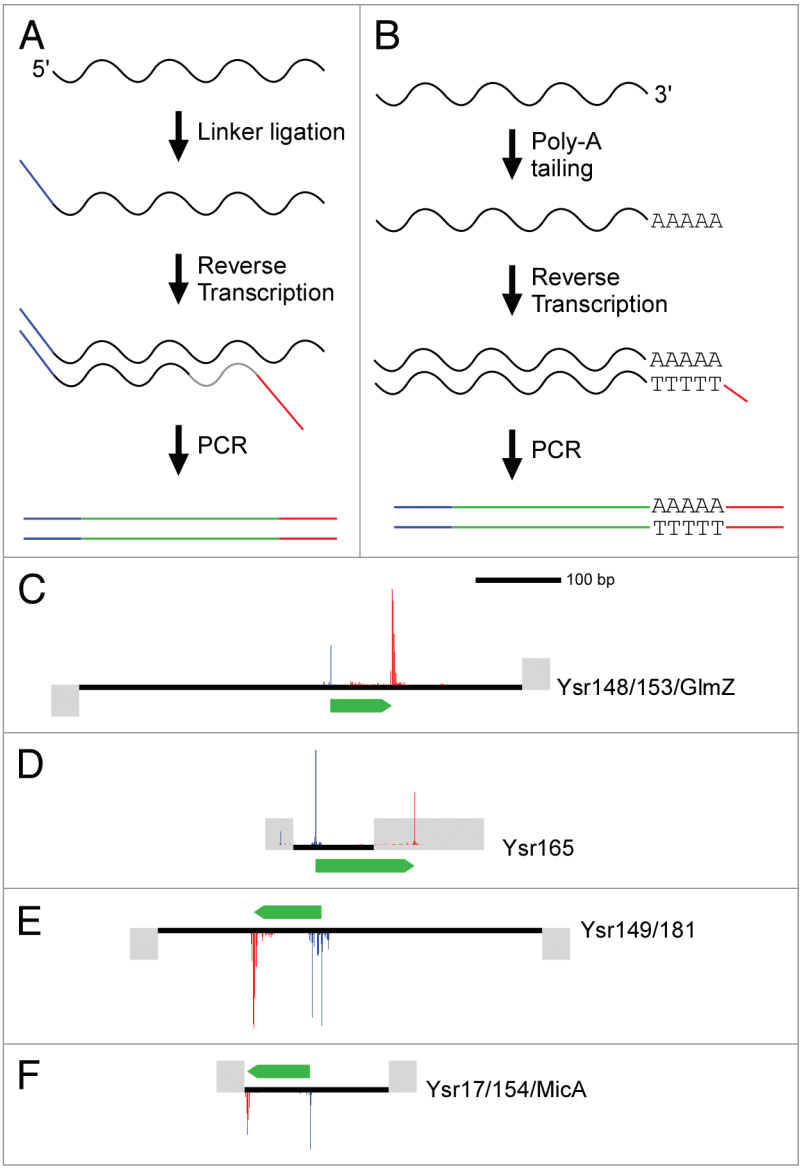

Figure 2. Representative examples of 5' and 3' Deep RACE data. (**A**) Schematic for the Deep 5' RACE method. (**B**) Schematic for the Deep 3' RACE method. (**C**) Deep 5' RACE data (blue) and Deep 3' RACE data (green) for four selected sRNAs. Flanking annotated genes are indicated by gray boxes, which are positioned above (plus) or below (minus) the horizontal line to denote strand.

Sequence conservation of sRNAs between *Y. pestis*, *Y. pseudotuberculosis*, *Y. enterocolitica* and *E. coli*. We used BLAST to search for sequence conservation between each of the sRNAs and the genomes of *Y. pseudotuberculosis*, *Y. enterocolitica* and *E. coli*. Specifically, we searched using the sequence beginning 100 bp upstream and ending 100 bp downstream of the coordinates identified in the initial deep sequencing experiment. A summary of this analysis is shown in **Table 2**. For sequences with matches in any of these species, we performed alignments using ClustalW (**Fig. S3**). In all cases, sRNA sequences in *Y. pestis* and *Y. pseudotuberculosis* were extremely similar, as expected due to the high sequence identity between these species. In most cases, nucleotide identity was ≥ 99%, including for Ysr172 which is not detectably expressed in *Y. pseudotuberculosis* (**Fig. 1**). All but two of the sRNAs are conserved in *Y. enterocolitica*, suggesting that they represent a core set of sRNAs for this genus. Only 14 of the sRNAs are conserved in *E. coli*. Given that most of the *E. coli* homologs have been characterized, conservation of sRNAs between *E. coli* and *Y. pestis* provides insight into the function of these sRNAs in *Y. pestis*. Several were only partially conserved, suggesting that their functions have diverged between the two species. Three *Y. pestis* sRNAs that did not generate a BLAST match in *E. coli* are located in the same gene context (i.e., same synteny with flanking genes) as known *E. coli* sRNAs, e.g., Ysr185/Spot 42. We propose that these sRNAs are shared between the two species but have diverged extensively with respect to their mRNA targets. As described previously for MicF (one of the 18 putative sRNAs that failed the northern blot analysis), some sRNAs conserved between *Y. pestis* and *E. coli* are conserved only over short stretches of sequence that are known to be required for base-pairing with targets in *E. coli*.[28]

Table 2. Summary of alignments of nucleotide sequences of identified RNAs

	Y. pseudotuberculosis[b]		*Y. enterocolitica*		*E. coli*	
sRNA[a]	**% identity**[c]	**% coverage**[d]	**% identity**[c]	**% coverage**[d]	**% identity**[c]	**% coverage**[d]
Ysr151/RnpB	99	100	92	100	89	85
Ysr88/152	100	100	83	64	N/A	N/A
Ysr148/153/GlmZ	100	79	72	61	N/A	N/A
Ysr7/154/MicA	99	100	86	99	80	33
Ysr155/RyfD	100	100	91	100	79	93
Ysr156/Ffs	99	100	92	72	84	40
Ysr145/157	99	100	79	100	N/A	N/A
Ysr158	99	100	69	40	N/A	N/A
Ysr159/CyaR	99	100	83	73	N/A	N/A
Ysr23/160	98	100	73	52	N/A	N/A
Ysr161	93	100	82	55	N/A	N/A
Ysr163	82	89	N/A	N/A	N/A	N/A
Ysr164	99	96	78	38	N/A	N/A
Ysr165	99	100	90	52	80	30
Ysr11/166/FnrS	100	100	87	57	83	41
Ysr167	99	100	79	82	N/A	N/A
Ysr73/169	99	100	76	100	N/A	N/A
Ysr170	100	100	N/A	N/A	N/A	N/A
Ysr171	100	100	84	69	79	42
Ysr172	100	78	87	22	83	25
Ysr173/*rmf*	100	76	89	75	74	72
Ysr174	100	100	82	54	81	25
Ysr65/175	100	66	93	36	N/A	N/A
Ysr177	99	100	79	100	N/A	N/A
Ysr179/CsrB	100	85	87	85	N/A	N/A
Ysr45/180/GcvB	100	100	94	75	80	75
Ysr149/181	100	100	91	68	N/A	N/A
Ysr182/6S RNA	99	100	89	79	78	78
Ysr183/SroG	99	100	92	58	70	52
Ysr185/Spot 42	96	100	76	98	N/A	N/A
Ysr186/CsrC	99	83	89	82	87	40
Ysr146.2/187	99	100	90	68	N/A	N/A

[a]Underlined RNAs were predicted in a bioinformatic study.[15] [b]The *Y. pseudotuberculosis* 32953 strain was used as a reference genome since no reference genome sequence exists for PTB52c. [c]% identity of sRNA between Y. pestis and representative organism. [d]% coverage of BLAST search of sRNA 100 bp upstream and downstream.

We propose that these sRNAs share a "core" set of mRNA targets across the *Enterobacteriaceae* that rely on the conserved sequence for base-pairing, but also have species-specific mRNA targets. In a few cases, e.g., Ysr165, Ysr172, sequence identity with *E. coli* was found only for the sequence flanking the sRNA ends. Thus, the sequence similarity is unlikely to reflect functional conservation of the sRNA.

Given that almost all of the sRNAs and their flanking sequence are > 95% identical between *Y. pestis* and *Y. pseudotuberculosis*, what accounts for the differences in expression patterns? One possibility is that Hfq functions differently in the two species. Hfq has been implicated previously in promoting stability of many sRNAs in other bacterial species.[7] However, Hfq is 100% identical at the amino acid sequence level between *Y. pestis* and *Y. pseudotuberculosis*, expression of *hfq* mRNA is similar between the two species (**Fig. 1A**) and the majority of sRNAs are 100% identical between the two species. Hence, it is unlikely that differential binding of Hfq alone contributes to the differences in expression patterns. Transcription of the sRNAs may be regulated differently between the two species, although any such differences would be due to *trans*-acting factors since the sequences upstream of the sRNAs are also well-conserved between the two species. Our preferred model is that mRNA target availability determines the stability of the sRNAs. Thus, differences in the

abundance of mRNA targets for each sRNA between *Y. pestis* and *Y. pseudotuberculosis* would alter the dynamics of base-pairing, in turn, altering the susceptibility of sRNAs to degradation. Consistent with this hypothesis, pairing of sRNAs to their target mRNAs has been shown previously to promote sRNA degradation.[29] Differences in mRNA target abundance could also impact the size of the available pool of Hfq, which may be limiting.[30]

A previous study of sRNAs in *E. coli* identified likely homologs in *Y. pestis* based on sequence conservation.[15] Eight of these predictions are consistent with the sRNAs we identified (**Table 2**). However, we did not detect RNA-seq signal for several predicted sRNAs. This is likely due to the fact that the equivalent *E. coli* sRNAs are expressed under stress conditions, e.g., RprA expression is induced during stationary phase.[31] Therefore, we propose that many sRNAs were not detected in our study due to condition-specific expression patterns.

Comparison to another sRNA study in a *Yersinia* species. A recent study used deep sequencing to identify 150 putative sRNAs in *Y. pseudotuberculosis*.[23] Surprisingly, only 17 of the 31 confirmed sRNAs that we identified are shared with the list of putative sRNAs identified by Koo et al. Thus, there are substantial disparities between the two studies. We propose the following explanations for these disparities: (1) six of the putative sRNAs identified by Koo et al. are not conserved in *Y. pestis*; (2) some of the putative sRNAs identified by Koo et al. may have escaped detection in our study by virtue of generating an insufficient number of sequence reads in the initial deep sequencing analysis. This is likely to be the case for tmRNA, for which we detect > 100 sequence reads but fewer than 500, the cut-off we used. However, for many of the sRNAs identified by Koo et al., we detect few or no sequence reads in the corresponding location in *Y. pestis*; (3) in a few cases, putative sRNAs identified by Koo et al. were specific to a particular growth phase that differs from the conditions in our study. These sRNAs would likely be missed by our approach; (4) only 29 of the putative 150 sRNAs identified by Koo et al. were successfully validated by Northern analysis and/or RACE. Hence, it is possible that many of the remaining 121 candidates would be below our detection threshold by Northern analysis or are false positives. Consistent with this, 20 of the 49 putative sRNAs tested were undetectable by northern analysis.[23] Furthermore, the 49 putative sRNAs selected for northern analysis by Koo et al. have considerably higher expression levels, based on the deep sequencing data, than those that were not tested (median number of sequence counts ~7-fold higher for those tested by northern analysis). This greatly increases the likelihood that the untested putative sRNAs would be below the detection threshold of northern analysis or are false positives; (5) it is possible that there is a much larger pool of sRNAs and the two studies have each identified a subset of that pool. Consistent with this, deep sequencing has been used to identify > 500 putative sRNAs in *Vibrio cholerae*[16] and > 300 putative sRNAs in *E. coli*;[14] (6) Koo et al. identified sRNAs in *Y. pseudotuberculosis* whereas we identified sRNAs in *Y. pestis*. Although the DNA sequences corresponding to almost all these putative sRNAs are highly conserved between the two species, we have shown that sRNA expression patterns vary considerably between the two species. Hence, it is likely that many of the putative sRNAs identified by Koo et al. are well expressed in *Y. pseudotuberculosis* but would be below the limit of detection, either by deep sequencing or Northern analysis, in *Y. pestis*.

Our northern blot data confirm the existence in *Y. pestis* of 13 of the 101 putative sRNAs identified by Koo et al. but not tested by northern blot in that study. Furthermore, our data provide the first characterization of the expression of these 13 sRNAs in *Y. pestis* and *Y. pseudotuberculosis*. The 50 putative sRNAs we identified in *Y. pestis* by deep sequencing include relatively few false positives, as evidenced by the high success rate when testing by northern blot. This is likely due to the stringent cut-off used for assignment as a putative sRNA. In contrast, the 150 putative sRNAs identified by Koo et al. likely include many false positives. It is also likely that the list of putative sRNAs identified by Koo et al. has many fewer false negatives than our list. By careful comparison of the data sets from both studies, it may be possible to prioritize additional putative sRNAs for validation.

Materials and Methods

Strains and growth conditions. Strains of *Y. pestis* used in this study were KIM6+ (Pgm$^+$ pCD1$^-$ pMT1$^+$ pPst$^+$),[32] KIM6+ Δ*hfq:cat*,[12] KIM6+ Δ*hfq*-multi-copy complementation:*kan* (Δ*hfq*-C),[12] CO92[33] and CO92 Δ*hfq:cat*.[12] Strains of *Y. pseudotuberculosis* used were PTB52c WT (pYV$^-$; serotype IB; YP-HPI$^+$),[34] PTB52c Δ*hfq:cat*,[12] PTB52c Δ*hfq*-multi-copy complementation:*kan* (Δ*hfq*-C),[12] PTB51c (pYV$^-$; serotype IB; YP-HPI$^+$),[34] PTB57c (pYV$^-$; serotype III; YP-HPI$^-$) and PTB54c (pYV$^-$; serotype III; YP-HPI$^-$).[34] Construction of mutant strains and growth conditions used in this study have been described previously.[12] All strains of *Y. pestis* and *Y. pseudotuberculosis* were grown in brain heart infusion (BHI) media. To obtain cells for RNA isolation, 5 ml of BHI was diluted 1:5 with overnight culture and cells were grown for 4 h at 28°C or 37°C. Cells were then harvested at 4°C and stored at -80°C.

RNA isolation and purification. Cells were resuspended in 1 ml TRIzol (Invitrogen), incubated at room temperature for 5 min and centrifuged at 12,000 × g for 10 min at 4°C (all subsequent centrifugations were performed at this temperature). The supernatant was transferred to a new microfuge tube and 200 μl chloroform:isoamyl alcohol (24:1 ratio) was added. The sample was shaken vigorously for 15 sec, incubated at room temperature for 3 min and centrifuged at 12,000 × g for 15 min. The aqueous phase was transferred to a new microfuge tube where 500 μl isopropanol was added and incubated for 10 min at room temperature. Following centrifugation at 12,000 × g for 10 min, the supernatant was decanted and washed with 1 ml ice-cold 75% ethanol and centrifuged at 7,600 × g for 5 min. The supernatant was decanted, and the residual supernatant was removed by pipette. The RNA pellet was air-dried, then resuspended in 30 μl of RNase-free water.

Resulting RNA was treated by DNase I (New England Biolabs) to remove any remaining DNA. A total of 10 μl DNase I was utilized in a final volume of 500 μl and incubated for 1 h at 37°C at which point 600 μl of isopropanol was added and precipitated overnight at -80°C. Following precipitation, the RNA was centrifuged at 12,000 × g for 20 min. The supernatant was decanted and washed with 1 ml ice-cold 75% ethanol and centrifuged

at 7,600 × g for 5 min. The supernatant was decanted, and the residual supernatant was removed by pipet. The RNA pellet was air-dried, then resuspended in 30 μl of RNase-free water. RNA was used for initial deep sequencing screening and replicate RNA samples were pooled, quantitated and aliquoted into microfuge tubes for use in Northern analysis.

Initial deep sequencing and characterization of sRNAs. Isolated RNA from *Y. pestis* KIM6+ grown at 37°C was separated on a 6% denaturing polyacrylamide gel and RNA below 400 nt was cut from the gel. The resulting RNA was electro-eluted using dialysis tubing following washing with 1 ml of 0.1 X TBE. Electro-elution was run at 100 V for 30 min. The resulting TBE was ethanol precipitated and the resulting RNA was examined on an agarose gel to ensure there was no residual rRNA. A cDNA library was constructed using the Illumina RNA-seq kit, following the manufacturer's instructions except that DNA was gel-purified from 200 bp and above rather than 300 bp before the PCR amplification step. This modification increased the likelihood of identifying sRNAs.

The DNA library was sequenced using an Illumina Genome Analyzer II (Harvard Medical School). Reads were mapped to the *Y. pestis* KIM genome using Bowtie with default settings.[35] Sequences were piled up to determine the number of sequence reads that mapped to each nucleotide of the genome. Putative sRNAs were identified as regions of contiguous sequence that partially or fully overlap an intergenic region (JCVI genome annotation) with at least one position with > 500 mapped sequence reads. Fully antisense RNAs could not be identified due to the lack of strand information in the sequencing data.

Northern transfer and hybridization. For Northern analysis, a total of 15 μg of RNA was separated on a 1.5% formaldehyde MOPS gel and transferred to a nylon membrane by capillary action. Hybridization was performed using 60-mer oligonucleotides (**Table S3**) that were γ^{32}P-ATP end-labeled with T4 polynucleotide kinase (Fermentas) for 1 h at 37°C. Membranes were hybridized at 42°C for 2 h in Amersham Rapid-hyb Buffer (GE Life Sciences) and were washed as per manufacturer's protocol. Densitometric quantitation of Northern Blots was performed using ImageQuant software and Sum Above Background calculations. Percent values indicated in **Figure 1B** are normalized to the condition with the highest signal. Four blots were reprobed with an oligonucleotide specific to 5S rRNA, as a loading control (**Fig. S6**).

Deep 5' and 3' RACE and computational analysis. RACE experiments were performed using RNA from *Y. pestis* KIM6+ RNA isolated at 37°C and the FirstChoice® RLM-RACE Kit (Ambion). For 5' RACE, we used a modified 5' RLM-RACE Protocol. A total of 8 μg RNA was treated with tobacco acid pyrophosphatase (TAP) at 37°C for 1 h, followed by ligation of the 5' RACE adaptor 37°C for 1 h. The resulting RNA was then reverse transcribed according to the manufacturer's protocol. PCR was then performed on the resulting cDNA using a primer containing the Ion Torrent sequence for 5' RACE amplification and primers specific for each sRNA confirmed by northern analysis (**Table S4**). For some sRNAs, the PCR had to be re-amplified using a portion of the original reaction and the same primers, which then resulted in successful amplification of bands for all sRNAs. Resulting PCR products were purified using the QIAquick PCR Purification Kit and eluted in "Low TE" (10 mM Tris, 0.1 mM EDTA). All products were then pooled together and sent for Ion Torrent deep sequencing using a 314 chip (Wadsworth Center Applied Genomic Technologies Core Facility).

For 3' RACE, we utilized the miScript Reverse Transcription Kit (Qiagen) to perform reverse transcription on *Y. pestis* KIM6+ RNA isolated at 37°C according to the manufacturer's protocol. PCR was performed on the resulting cDNA using the primers containing Ion Torrent sequence for 3' RACE (universal primer) and each sRNA. Resulting PCR products were purified using the QIAquick PCR Purification Kit, eluted in "Low TE," pooled and sent for Ion Torrent deep sequencing using a 314 chip (Wadsworth Center Applied Genomic Technologies Core Facility).

Any sequences lacking the expected 5' or 3' adaptor sequences were removed. We then extracted non-adaptor sequence from the remaining reads and mapped them to the *Y. pestis* KIM genome using BWA with default parameters.[36] Reads were assumed to be associated with an sRNA if they were located within 1 kbp of the predicted location and were located on the predicted strand.

Clustering analysis of Ysrs verified by northern blot analysis. To assign expression patterns of Ysrs into groups, we used *k*-means clustering to partition the sRNAs into $k = 9$ clusters, selecting k by minimizing the value of the Kelley penalty.[37] These nine groups were manually adjusted to seven for clarity.

Conservation analysis. Sequences from 100 bp upstream to 100 bp downstream of each sRNA (coordinates derived from the original Illumina sequencing data) were used to search against the *Y. pseudotuberculosis* 32953, *Y. enterocolitica* 8081 and *E. coli* K-12 (MG1655) strains using BLAST with the default parameters.[38] BLAST matches were realigned using ClustalW.[39]

Conclusions

In summary, we have identified 32 sRNAs in *Y. pestis*, of which 14 are novel, and 11 of these 14 have no known *E. coli* homolog. Based on the patterns of sRNA expression and on the differences between sRNA expression in *Y. pestis* and *Y. pseudotuberculosis*, we propose that many of these sRNAs contribute to the unique biology of *Y. pestis*, and may play important roles in virulence.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank the Applied Genomic Technologies Core Facility for technical assistance. We thank Keith Derbyshire, Todd Gray, Kris Spaeth, David Grainger and members of the Wade group for helpful discussions. This work was funded by the National Institutes of Health through the NIH Director's New Innovator Award Program, 1DP2OD007188 (J.T.W.) and by Wadsworth Center "SIG" funds to J.T.W. and K.A.M.

Supplemental Materials

Supplemental material may be found here:
www.landesbioscience.com/journals/rnabiology/article/23590

References

1. Achtman M, Zurth K, Morelli G, Torrea G, Guiyoule A, Carniel E. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis.* Proc Natl Acad Sci USA 1999; 96:14043-8; PMID:10570195; http://dx.doi.org/10.1073/pnas.96.24.14043.
2. Skurnik M, Peippo A, Ervelä E. Characterization of the O-antigen gene clusters of *Yersinia pseudotuberculosis* and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. Mol Microbiol 2000; 37:316-30; PMID:10931327; http://dx.doi.org/10.1046/j.1365-2958.2000.01993.x.
3. Chain PS, Carniel E, Larimer FW, Lamerdin J, Stoutland PO, Regala WM, et al. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis.* Proc Natl Acad Sci USA 2004; 101:13826-31; PMID:15358858; http://dx.doi.org/10.1073/pnas.0404012101.
4. Perry RD, Fetherston JD. *Yersinia pestis*--etiologic agent of plague. Clin Microbiol Rev 1997; 10:35-66; PMID:8993858.
5. Mayer L, Greenstein AJ. Acute yersinial ileitis: a distinct entity. Am J Gastroenterol 1976; 65:548-51; PMID:961687.
6. Waters LS, Storz G. Regulatory RNAs in bacteria. Cell 2009; 136:615-28; PMID:19239884; http://dx.doi.org/10.1016/j.cell.2009.01.043.
7. Vogel J, Luisi BF. Hfq and its constellation of RNA. Nat Rev Microbiol 2011; 9:578-89; PMID:21760622; http://dx.doi.org/10.1038/nrmicro2615.
8. Chao Y, Vogel J. The role of Hfq in bacterial pathogens. Curr Opin Microbiol 2010; 13:24-33; PMID:20080057; http://dx.doi.org/10.1016/j.mib.2010.01.001.
9. Geng J, Song Y, Yang L, Feng Y, Qiu Y, Li G, et al. Involvement of the post-transcriptional regulator Hfq in *Yersinia pestis* virulence. PLoS One 2009; 4:e6213; PMID:19593436; http://dx.doi.org/10.1371/journal.pone.0006213.
10. Schiano CA, Bellows LE, Lathem WW. The small RNA chaperone Hfq is required for the virulence of *Yersinia pseudotuberculosis.* Infect Immun 2010; 78:2034-44; PMID:20231416; http://dx.doi.org/10.1128/IAI.01046-09.
11. Rempe KA, Hinz AK, Vadyvaloo V. Hfq regulates biofilm gut blockage that facilitates flea-borne transmission of *Yersinia pestis.* J Bacteriol 2012; 194:2036-40; PMID:22328669; http://dx.doi.org/10.1128/JB.06568-11.
12. Bai G, Golubov A, Smith EA, McDonough KA. The importance of the small RNA chaperone Hfq for growth of epidemic *Yersinia pestis*, but not *Yersinia pseudotuberculosis*, with implications for plague biology. J Bacteriol 2010; 192:4239-45; PMID:20543069; http://dx.doi.org/10.1128/JB.00504-10.
13. Raghavan R, Groisman EA, Ochman H. Genome-wide detection of novel regulatory RNAs in *E. coli.* Genome Res 2011; 21:1487-97; PMID:21665928; http://dx.doi.org/10.1101/gr.119370.110.
14. Shinhara A, Matsui M, Hiraoka K, Nomura W, Hirano R, Nakahigashi K, et al. Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli.* BMC Genomics 2011; 12:428; PMID:21864382; http://dx.doi.org/10.1186/1471-2164-12-428.
15. Hershberg R, Altuvia S, Margalit H. A survey of small RNA-encoding genes in *Escherichia coli.* Nucleic Acids Res 2003; 31:1813-20; PMID:12654996; http://dx.doi.org/10.1093/nar/gkg297.
16. Liu JM, Livny J, Lawrence MS, Kimball MD, Waldor MK, Camilli A. Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. Nucleic Acids Res 2009; 37:e46; PMID:19223322; http://dx.doi.org/10.1093/nar/gkp080.
17. Irnov I, Sharma CM, Vogel J, Winkler WC. Identification of regulatory RNAs in *Bacillus subtilis.* Nucleic Acids Res 2010; 38:6637-51; PMID:20525796; http://dx.doi.org/10.1093/nar/gkq454.
18. Vockenhuber MP, Sharma CM, Statt MG, Schmidt D, Xu Z, Dietrich S, et al. Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor.* RNA Biol 2011; 8:468-77; PMID:21521948; http://dx.doi.org/10.4161/rna.8.3.14421.
19. Albrecht M, Sharma CM, Dittrich MT, Müller T, Reinhardt R, Vogel J, et al. The transcriptional landscape of *Chlamydia pneumoniae.* Genome Biol 2011; 12:R98; PMID:21989159; http://dx.doi.org/10.1186/gb-2011-12-10-r98.
20. Mitschke J, Georg J, Scholz I, Sharma CM, Dienst D, Bantscheff J, et al. An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis sp*. PCC6803. Proc Natl Acad Sci USA 2011; 108:2124-9; PMID:21245330; http://dx.doi.org/10.1073/pnas.1015154108.
21. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori.* Nature 2010; 464:250-5; PMID:20164839; http://dx.doi.org/10.1038/nature08756.
22. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, et al. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. PLoS Genet 2008; 4:e1000163; PMID:18725932; http://dx.doi.org/10.1371/journal.pgen.1000163.
23. Koo JT, Alleyne TM, Schiano CA, Jafari N, Lathem WW. Global discovery of small RNAs in *Yersinia pseudotuberculosis* identifies *Yersinia*-specific small, noncoding RNAs required for virulence. Proc Natl Acad Sci USA 2011; 108:E709-17; PMID:21876162; http://dx.doi.org/10.1073/pnas.1101655108.
24. Olivarius S, Plessy C, Carninci P. High-throughput verification of transcriptional starting sites by Deep-RACE. Biotechniques 2009; 46:130-2; PMID:19317658; http://dx.doi.org/10.2144/000113066.
25. Thomason MK, Storz G. Bacterial antisense RNAs: how many are there, and what are they doing? Annu Rev Genet 2010; 44:167-88; PMID:20707673; http://dx.doi.org/10.1146/annurev-genet-102209-163523.
26. Wada A, Yamazaki Y, Fujita N, Ishihama A. Structure and probable genetic location of a "ribosome modulation factor" associated with 100S ribosomes in stationary-phase *Escherichia coli* cells. Proc Natl Acad Sci USA 1990; 87:2657-61; PMID:2181444; http://dx.doi.org/10.1073/pnas.87.7.2657.
27. Hobbs EC, Fontaine F, Yin X, Storz G. An expanding universe of small proteins. Curr Opin Microbiol 2011; 14:167-73; PMID:21342783; http://dx.doi.org/10.1016/j.mib.2011.01.007.
28. Delihas N. Annotation and evolutionary relationships of a small regulatory RNA gene *micF* and its target *ompF* in *Yersinia* species. BMC Microbiol 2003; 3:13; PMID:12834539; http://dx.doi.org/10.1186/1471-2180-3-13.
29. Massé E, Escorcia FE, Gottesman S. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli.* Genes Dev 2003; 17:2374-83; PMID:12975324; http://dx.doi.org/10.1101/gad.1127103.
30. Hussein R, Lim HN. Disruption of small RNA signaling caused by competition for Hfq. Proc Natl Acad Sci USA 2011; 108:1110-5; PMID:21189298; http://dx.doi.org/10.1073/pnas.1010082108.
31. Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15:1637-51; PMID:11445539; http://dx.doi.org/10.1101/gad.901001.
32. Deng W, Burland V, Plunkett G 3rd, Boutin A, Mayhew GF, Liss P, et al. Genome sequence of Yersinia pestis KIM. J Bacteriol 2002; 184:4601-11; PMID:12142430; http://dx.doi.org/10.1128/JB.184.16.4601-4611.2002.
33. Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MT, Prentice MB, et al. Genome sequence of Yersinia pestis, the causative agent of plague. Nature 2001; 413:523-7; PMID:11586360; http://dx.doi.org/10.1038/35097083.
34. Hare JM, McDonough KA. High-frequency RecA-dependent and -independent mechanisms of Congo red binding mutations in *Yersinia pestis.* J Bacteriol 1999; 181:4896-904; PMID:10438760.
35. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009; 10:R25; PMID:19261174; http://dx.doi.org/10.1186/gb-2009-10-3-r25.
36. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009; 25:1754-60; PMID:19451168; http://dx.doi.org/10.1093/bioinformatics/btp324.
37. Kelley LA, Gardner SP, Sutcliffe MJ. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Eng 1996; 9:1063-5; PMID:8961360; http://dx.doi.org/10.1093/protein/9.11.1063.
38. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990; 215:403-10; PMID:2231712.
39. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22:4673-80; PMID:7984417; http://dx.doi.org/10.1093/nar/22.22.4673.