Virulence 3:2, 154–156; March/April 2012; © 2012 Landes Bioscience

Identification of small, noncoding RNAs in pathogenic *Yersinia* species

Implications for evolution and virulence

Wyndham W. Lathem

Department of Microbiology-Immunology; Northwestern University Feinberg School of Medicine; Chicago, IL USA

Small, noncoding RNAs (sRNAs) are major regulators of gene expression in bacteria, affecting protein synthesis at the post-transcriptional level. We recently reported the identification of the global set of sRNA molecules expressed by the enteric pathogen *Yersinia pseudotuberculosis*. This bacterium causes the disease Yersiniosis and is considered the evolutionary ancestor of the closely related plague pathogen *Yersinia pestis*. We discovered that approximately 80% of the *Y. pseudotuberculosis* "sRNA-ome" is specific for *Yersinia* species and comparisons with *Y. pestis* identified multiple sRNA genes unique to *Y. pseudotuberculosis*. We also found differences between the species in both sRNA expression and dependence on the sRNA chaperone Hfq. Here we suggest that these differences may have contributed to the unique disease etiologies of *Y. pestis* and *Y. pseudotuberculosis* through evolutionary changes in the post-transcriptional regulatory networks that are required for virulence.

Keywords: sRNA, Hfq, post-transcriptional regulation, plague, Yersiniosis

Submitted: 11/30/11

Revised: 12/21/11

Accepted: 12/22/11

http://dx.doi.org/10.4161/viru.19155

Correspondence to: Wyndham W. Lathem; Email: lathem@northwestern.edu

News and View to: Koo JT, Alleyne TM, Schaino 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

The ability of bacteria to cause disease depends not only on the repertoire of virulence factors that mediate the host-pathogen interaction but also on the proper regulation of those factors during the various stages of infection. When one considers the flow of information in the cell, there are multiple points at which gene expression and protein synthesis may be controlled. These include (but are not limited to) alterations in gene copy number (e.g., single copy on the chromosome vs. multi-copy on plasmids); transcriptional regulation through sigma factors, activators and repressors, enhancers, and DNA conformation; post-transcriptional regulation via mRNA stability, efficiency of ribosome binding, codon usage, riboswitches, and *cis*- and *trans*-acting noncoding RNA molecules; and post-translational regulation due to protein stability or turnover, structure, modifications, and requirements for co-factors (e.g., cyclic diguanylate or cAMP). Coordination of each of these regulatory steps requires a precise architectural program that results in optimal virulence factor expression. In addition, fine-tuning of multiple regulatory pathways that takes into account the environmental niche, nutritional status, and stress conditions affecting the cell is necessary for the fitness of the pathogen and induction of the disease state.

One mechanism by which this fine-tuning occurs is through the post-transcriptional regulation of protein synthesis mediated by the expression and activity of small, noncoding RNA (sRNA) molecules. Bacterial sRNAs interact with target mRNA sequences through short (6–8 nucleotides), imperfect base-pairing; this interaction often results in changes in mRNA stability and/or efficiency of translation. Usually the consequence of the sRNA-mRNA interaction is the downregulation of protein synthesis, but in some cases increased translation may result. For more detail on the mechanisms by which sRNAs function, please see the review by Gottesman and Storz, published in 2010. Most sRNAs are encoded in the intergenic regions of the genome and contain their own promoters and rho-independent terminators. Once transcribed, some sRNAs are known to interact with proteins as well as mRNAs. For example, the sRNAs CsrB and CsrC bind to and act as antagonists of the carbon storage regulator protein CsrA, and the SsrA RNA forms a complex with

the protein chaperone SmpB to participate in *trans*-translation. In addition, a large number of trans-acting sRNAs are dependent on the small RNA chaperone protein Hfq for their expression, stability, and/or function. Hfq is thought to protect sRNAs from RNase degradation, accelerate RNA strand exchange and annealing, and stabilize the sRNA-mRNA interaction. Many bacterial pathogens require Hfq for their full virulence, demonstrating the importance of sRNA function during the pathogenic process.

Indeed, our laboratory recently reported that Hfq is critical for the full virulence of the gastrointestinal pathogen *Yersinia pseudotuberculosis* (Schiano et al. 2010). Transmitted by the ingestion of contaminated food or water, *Y. pseudotuberculosis* is primarily a pathogen of wild animals and birds. In humans, *Y. pseudotuberculosis* causes the mild, self-limiting disease Yersiniosis and is characterized by ileitis, mesenteric lymphadenitis, fever, and diarrhea. Individuals usually develop symptoms 5–10 d following ingestion of bacteria, and the disease typically lasts 1–3 weeks in the absence of antibiotic treatment. In immunocompromised patients, systemic dissemination of *Y. pseudotuberculosis* may occur that can result in a potentially life-threatening infection, although this is rare. *Y. pseudotuberculosis* possesses a formidable armament of virulence factors to cause disease in humans, including the adhesins invasin, YadA, and Ail, the superantigen YPM, the Yersiniabactin iron acquisition system, and the pYV-encoded type III secretion system (T3SS). The T3SS delivers a series of effector proteins called Yops directly into the cytosol of host cells that disrupts the host cell cytoskeleton, prevents phagocytosis of the bacterium by immune cells, and inhibits the production of pro-inflammatory cytokines. This T3SS is essential for the virulence of all *Yersinia* species pathogenic to humans.

We found that the loss of Hfq from *Y. pseudotuberculosis* affects multiple virulence-associated attributes of the bacterium, including motility, resistance to oxidative stress, adherence to eukaryotic cells and intracellular survival, and the production and secretion of multiple T3SS Yop effector proteins. The consequence of this disregulation is a 10,000-fold attenuation of *Y. pseudotuberculosis* in a mouse model of Yersiniosis and suggests that one or more Hfq-dependent sRNAs participates in the proper control of Yersinia virulence factor expression during mammalian infection. In order to understand how sRNAs contribute to the optimal regulation of virulence in this enteropathogen, we surveyed the literature but found only 15 noncoding RNAs annotated in the genome sequence of *Y. pseudotuberculosis*. As a computational analysis by Livny et al. (2006) predicted that ~1,400 sRNAs may be encoded within the genome of the closely related species *Y. pestis*, and other Gram-negative species have been reported to express 100 or more sRNAs, it seemed likely that the *Y. pseudotuberculosis* genome annotation for sRNAs was incomplete. Therefore, we set out to identify the global set of sRNAs expressed by this *Yersinia* species at four different points along the growth curve (early-, mid-, late-log and stationary phases) at both 26°C and 37°C, two physiologically relevant temperatures for *Y. pseudotuberculosis*.

To accomplish our goal, we sequenced a size-selected transcriptome of *Y. pseudotuberculosis* using next generation, high-throughput technology based on the Illumina-Solexa platform. This study resulted in up to 17 million sequencing reads per sample, and based on homology to the *Y. pseudotuberculosis* IP32953 genome, each read was classified as mRNA, tRNA, rRNA, "miscellaneous" RNA, or transcribed from the intergenic regions of the genome. As most trans-encoded sRNA genes are found within these intergenic regions, we applied a filtering algorithm based on the characteristic features of sRNAs (including the presence of predicted promoters and rho-independent terminators) to eliminate 5' or 3' mRNA untranslated regions from the intergenic sequences. This analysis confirmed the expression of all 15 previously annotated "miscellaneous" noncoding RNAs such as MicF, spf, CsrC and SsrA, demonstrating the effectiveness of our approach to identify sRNAs by deep sequencing. In addition, we identified 150 previously unannotated sRNAs expressed by *Y. pseudotuberculosis*, which we refer to as Ysrs (for Yersinia small RNAs).

While 32 of these previously unannotated sRNAs are represented by orthologous sequences in the *Escherichia coli* and *Salmonella typhimurium* genomes, 79% of the Ysr genes we identified are specific to *Yersinia* species in that they do not show sequence conservation with other bacteria. In addition, numerous sRNAs present in other enteric species are missing from *Y. pseudotuberculosis*, including ArcZ, DicF, DsrA, CyaR and others. This suggests that the sRNA-ome of Yersinia has adapted to the particular lifestyle and niche of this organism, both gaining sRNA genes that are required for its survival in the mammalian host or in the environment, and losing those sRNAs that are dispensable or potentially detrimental.

With this in mind, we examined whether any of the Yersinia-specific sRNAs we identified in this study contribute to the virulence of *Y. pseudotuberculosis* in a mouse model of Yersiniosis. We generated isogenic mutants deleted for a number of sRNA genes and found that the loss of at least two sRNAs, Ysr29 and Ysr35, resulted in a significant attenuation of *Y. pseudotuberculosis*. Our results suggest that these sRNAs regulate factors that are required for virulence (be it canonical virulence factors, transporters, regulators, stress response proteins, etc.), and as Ysr29 and Ysr35 are Yersinia-specific, that the pathogen has adapted post-transcriptional regulatory networks controlled by these sRNAs to optimize the survival of *Y. pseudotuberculosis* during infection. Interestingly, the loss of RybB, a conserved sRNA encoded in the genomes of numerous bacteria that regulates the synthesis of multiple outer membrane proteins (OMPs), had no impact on the virulence of *Y. pseudotuberculosis* in this model. It is unclear whether the RybB regulon includes OMPs required for virulence or even if RybB is expressed by *Y. pseudotuberculosis* during mammalian infection. It remains to be determined if RybB contributes to the virulence of other pathogens such as diarrheagenic *E. coli* or Salmonella, and if so, how the RybB regulon between these species overlaps. In turn, this may inform studies on the evolution of enteric pathogens, the niches they occupy, and the factors they require to cause disease.

Although our comparison of the *Y. pseudotuberculosis* sRNA-ome with that of other enteric species revealed only 20% overlap in orthologous sRNA genes, a similar analysis of the *Yersinia pestis* CO92 genome found 96% of the unannotated sRNA gene sequences are shared. While infection with *Y. pseudotuberculosis* may result in gastrointestinal illness, *Y. pestis* is the causative agent of plague, an acute, rapidly progressing, febrile disease that is transmitted between animals by infected insect vectors (typically fleas) or through the inhalation of respiratory droplets carrying the bacteria. There are three distinct forms of the disease: bubonic, pneumonic, and septicemic, and in all cases, early treatment with antibiotics is crucial, otherwise mortality rates can reach 100% (Perry and Fetherston, 1997). Even though the routes of transmission, clinical disease manifestations, and mortality rates caused by *Y. pestis* and *Y. pseudotuberculosis* are dramatically different, *Y. pestis* is considered an evolutionary clone of *Y. pseudotuberculosis*, having directly evolved from the enteropathogen somewhere between 1,500–20,000 y ago (Achtman et al. 1999).

As *Y. pseudotuberculosis* carries the genes for 6 sRNAs that are missing from *Y. pestis*—including Ysr29—our data suggest that changes in the sRNA-ome during the evolution of the plague pathogen may have also contributed to the unique disease manifestations of each species. In fact, while Ysr29 is specific to *Y. pseudotuberculosis*, all the proteins regulated by this sRNA, including outer membrane protein A, ribosomal protein S1, glutathione S-transferase, the chaperonin GroEL, urease, peroxidase, the chaperone DnaK, and ribosome recycling factor, are found in *Y. pestis*. This observation, then, raises the following questions: are these targets no longer regulated by sRNAs in *Y. pestis*? Or do alternative sRNAs regulate these factors in a functionally redundant manner?

Differences in the sRNA-ome between *Y. pestis* and *Y. pseudotuberculosis* are not just limited to the presence or absence of sRNA genes, however. We also found that both the expression pattern and dependence on Hfq of multiple sRNAs conserved between the species differs. In *Y. pestis*, the expression of several sRNAs peaks in late-log phase and decreases in stationary phase, but in *Y. pseudotuberculosis*, those same sRNAs are maximally expressed in stationary phase. In addition, in the plague pathogen the loss of Hfq leads to decreased stability and/or expression of sRNAs such as MicA, FnrS, and GcvB, but in *Y. pseudotuberculosis* these sRNAs do not require Hfq for their expression/stability. The reasons for these differences are not yet known but may involve alternate sRNA chaperones, regulators that control sRNA transcription, or changes in RNase activity.

Indeed, these differences may manifest in multiple ways. While Hfq is also required for the full virulence of *Y. pestis* during mammalian infection and contributes to intracellular survival and resistance to oxidative stress (Geng et al. 2010), the loss of Hfq from the plague bacillus, but not from *Y. pseudotuberculosis*, leads to a severe growth restriction at 37°C. In addition, several phenotypes observed in the Hfq mutant of *Y. pseudotuberculosis*, such as increases in non-flagellar-based motility and biosurfactant secretion, do not occur in Δ*hfq* strains of *Y. pestis*. As the evolution of *Y. pestis* altered its routes of transmission and disease etiologies, it is likely that the expression of some sRNAs increased in importance while the reliance on others diminished to adapt to the changing environmental niches and to respond to different regulatory stimuli. Both the gain and loss of plasmid-based virulence factors such as the F1 antigen, Pla, and YadA may have also shifted the global sRNA regulon to increase core and accessory genome cross-talk mediated via sRNA-dependent regulatory networks, as has recently been demonstrated in Salmonella by Pfeiffer et al. (2007). It will be of interest to determine if any sRNAs in either species enables optimal expression of virulence characteristics that are specific for one or the other bacterium, and how those networks have changed (if at all) in the related species.

Evolution of species-specific virulence characteristics or host-microbe interaction dynamics through changes in regulatory networks is well established. For instance, Mandel et al. (2009) demonstrated that the ability of *Vibrio fischeri* ES114 to colonize the bobtail squid *Euprymna scolopes* relies on the two-component sensor kinase RscS, which allows *V. fischeri* to form biofilms through the induction of the Syp exopolysaccharide biosynthetic locus. The fish symbiont *V. fischeri* strain MJ11 lacks the *rscS* gene, rendering it unable to colonize the bobtail squid even though it contains the Syp locus. Similarly, while both *Y. pestis* and *Y. pseudotuberculosis* share the structural genes required for biofilm formation, only *Y. pestis* produces exopolysaccharides during insect colonization. Sun et al. (2008) determined that the acquisition of an inactivating mutation in the two-component regulatory accessory protein RcsA led to the derepression of biofilm formation that is required for vector-borne transmission. A functional RcsA protein in *Y. pseudotuberculosis* enhances RcsBCD-mediated inhibition of the biofilm synthesis genes; in *Y. pestis*, in the absence of RcsA this phosphorelay system is unable to prevent biofilm synthesis during flea colonization. Thus, changes in the regulation of biofilm synthesis, rather than the gain or loss of the virulence determinant itself, was a necessary step in the ability of *Y. pestis* to colonize and be transmitted by fleas, and thereby cause plague.

This study expands the hypothesis that bacterial speciation can occur through changes in gene regulation by including the evolution of post-transcriptional regulatory networks mediated by sRNAs. The identification of the sRNA-ome expressed by *Y. pseudotuberculosis* has not only discovered sRNAs specific to the Yersinia genus that are required for virulence, but has also revealed unexpected differences with the closely related plague pathogen *Y. pestis*. Further study on the conserved and distinct post-transcriptional regulatory networks found in related species should provide unique insights into the evolution of bacterial pathogens and will broaden our understanding of the global information flow in the cell.