Minireview

New levels of sophistication in the transcriptional landscape of bacteria

Arnoud HM van Vliet* and Brendan W Wren†

Addresses: *Institute of Food Research, Colney Lane, Norwich NR4 7UA, UK. †Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.

Correspondence: Brendan W Wren. Email: brendan.wren@lshtm.ac.uk

Abstract

An extra layer of complexity in the regulation of gene expression in bacteria is now apparent through previously unanticipated roles of noncoding and antisense RNAs.

Bacteria are the great survivors on planet Earth, where they can adapt and flourish in harsh environments ranging from deep-sea vents to acidic mine shafts. A feature of many bacteria, particularly pathogenic bacteria, is their ability to adapt and thrive in multiple environments, which provides them with a competitive advantage. For example, the facultative intracellular pathogen *Listeria monocytogenes* happily survives in the ambient environment as a saprophyte, but on occasions it has an inherent capacity to turn nasty and cause brain and materno-fetal infections in humans [1]. This requires the bacterium to switch genes on and off as it traverses different environments, ranging from a saprophytic lifestyle to the gut lumen after ingestion to invasion of epithelial cells and intracellular survival. The key to the survivalist success of pathogens is their ability to coordinate, redirect and fine-tune their genetic repertoire as and when required. Traditionally, transcriptional reshaping in bacteria has been considered to be controlled by a hierarchical network of interconnected global transcriptional regulators, such as sigma factors and one- and two-component regulatory systems [2]. In the past decade it has become apparent that the various forms of noncoding regulatory RNA (previously considered as intergenic junk) play important roles in the global regulation of cellular functions, and may represent connecting links between many cellular networks [3,4]. As such, noncoding RNA also plays a subtle but crucial role in the coordination of the expression of bacterial virulence determinants [5]. Two recent papers from Pascale Cossart and colleagues [6,7] present a comprehensive microarray analysis of the transcriptome of *Listeria monocytogenes* in different conditions, uncovering an unsuspected variety of regulatory roles for noncoding RNAs in controlling changes in gene expression that characterize the transition from saprophytic to pathogenic lifestyle.

Bacterial regulatory RNAs are more than intergenic junk

As well as the familiar types of RNA - messenger RNA, ribosomal RNA, and transfer RNA - bacteria also express many other noncoding RNAs (Figure 1). Some of these are catalytic, such as RNase P, or interact directly with proteins, like the 6S RNA that interacts directly with σ^{70}-containing RNA polymerase [8] or the CsrA-sequestering CsrB and CsrC RNAs [9]. Most bacterial noncoding RNAs, however, are thought to have roles in the post-transcriptional regulation of gene expression, using their capacity for complementary base-pairing [3,4]. Regulatory RNAs are usually subdivided into different groups in respect to their genomic position: one group contains those encoded *in cis* with their target gene (such as riboswitches and antisense RNAs); the other those encoded *in trans* from their target gene, which are often located at completely unrelated positions on the genome, such as the canonical small RNAs (sRNA).

Cis-encoded regulatory RNAs in principle enable a multitude of regulatory responses to stimuli, but they are mostly used for the sensing of temperature, metabolites, or metabolic stimuli (Figure 1a). They can, for example, function in translational control as riboswitches [4], where the full-length mRNA is preceded by a folded 5' untranslated region (5' UTR), which can fold in different conformations depending on the stimulus, often either allowing or blocking access to the Shine-Dalgarno ribosome-binding sequence, thus controlling translation (Figure 1a). Examples are the *prfA* thermosensor of *L. monocytogenes* [10], which is involved in controlling virulence genes, or the cyclic-di-GMP-sensing riboswitch of *Vibrio cholerae* [11], which controls biofilm formation, cell differentiation, and virulence gene expression. Alternatively, the 5' UTR can form transcriptional terminator or antiterminator loops (see Figure 1a), depending on the stimulus, and as such, control transcription of the downstream gene. An example of this is in the *Escherichia coli* tryptophan biosynthesis operon [12].
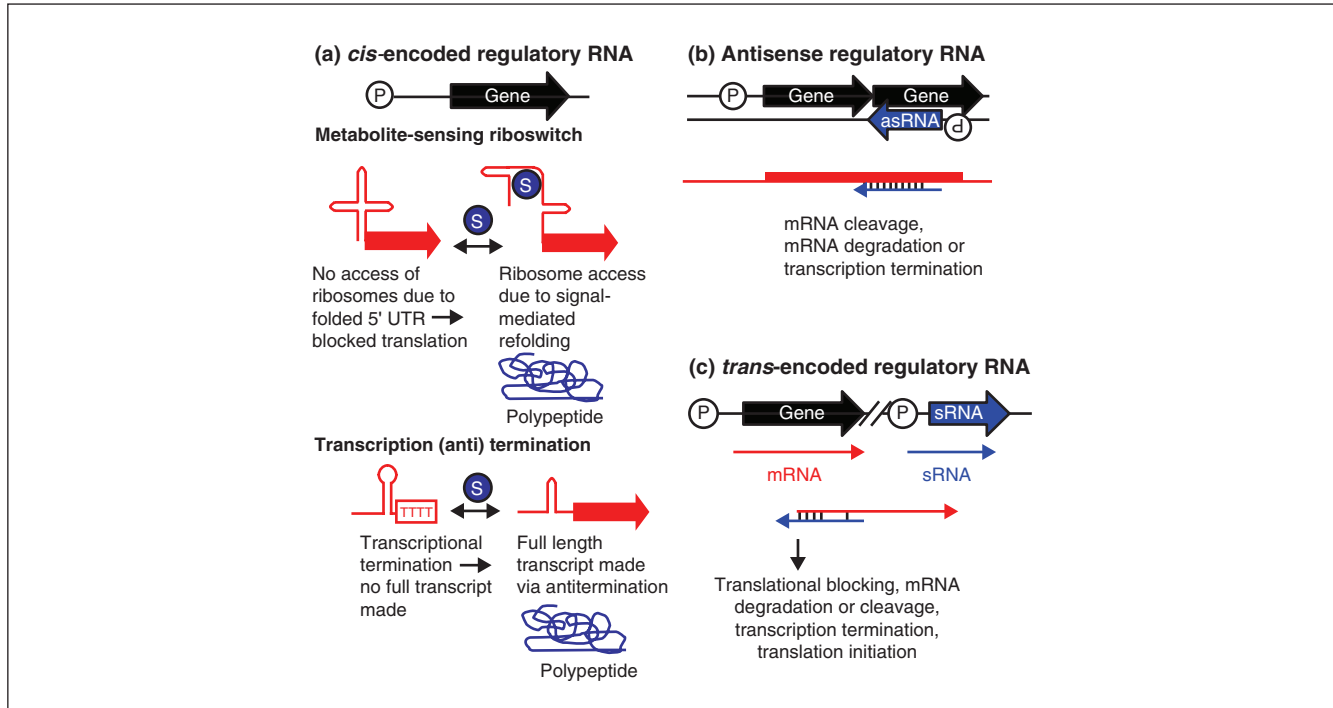
Figure 1

Schematic representation of the different modes of action of transcriptional regulation by noncoding RNA. **(a)** Regulation by *cis*-encoded RNA, subdivided into signal-mediated folding or unfolding of the translation initiation region (top) or transcription termination/antitermination (bottom). **(b)** Regulation by antisense RNA (asRNA). **(c)** Regulation by *trans*-encoded sRNA. Circled P, promoter; circled S, environmental or metabolic signal.

A relatively newly discovered class of *cis*-encoded regulatory RNA is antisense RNA (Figure 1b). Antisense RNA is known to be involved in bacterial Type I toxin-antitoxin systems [13] and the control of incompatibility of some plasmids [14], but has not been implicated in controlling the expression of virulence genes. However, this may be because microarrays used in functional genomics and transcriptome studies have not incorporated probes for potential antisense transcripts. Antisense RNA shows perfect complementarity to its target sequences, and can either silence gene expression or lead to alternative processing of the mRNA.

Trans-encoded regulatory RNAs are the best-known class of noncoding RNA (Figure 1c) and represent the prokaryotic version of RNA interference [15]. Bacterial sRNAs are anywhere between 30 and 500 nucleotides in length, but the complementary base-pairing is often restricted to a small aptamer of 8-20 nucleotides, which may have imperfect complementarity to its target mRNAs [3]. The effects of base-pairing range from translational blocking (when bound to the 5' UTR overlapping the Shine-Dalgarno sequence) to mRNA degradation, although there are also examples where sRNA binding induces translation of the target mRNA [3,16]. Most information on bacterial sRNA regulation has been obtained from analysis of *E. coli* and *Salmonella enterica* serovar Typhimurium, in which sRNA regulation controls many important aspects of cellular responses, ranging from outer membrane protein composition to metabolism [5].

Comprehensive transcriptome analysis and transcriptional reshaping

While the complexity of eukaryotic transcriptional regulation is well appreciated [17], it was long thought that bacteria lacked many of the sophisticated regulatory mechanisms affecting the cellular transcription landscape. However, recent papers have shed new light on this subject, and have revealed that notions of the relative simplicity of bacterial gene expression severely underestimates the potential of bacteria to control expression of their genetic repertoire [6,18-21].

Pascale Cossart and colleagues (Toledo-Arana *et al.* [6]) present one of the first comprehensive unbiased transcriptomes of a bacterium, in this case *L. monocytogenes*, that includes noncoding and antisense RNAs. This tour de force reveals that gene regulatory mechanisms in bacteria are far more complex than previously appreciated. Using an unbiased, high-density tiling microarray consisting of

345,668 25-base oligonucleotide probes covering the whole genome, they have shown antisense RNAs spanning several open reading frames and long overlapping 5' and 3' UTRs, in addition to 50 sRNAs. One surprising finding was the presence of long 5' UTRs which functioned as antisense RNA, as observed upstream of the *mogR* gene encoding a flagellar regulatory protein [6]. By comparing transcripts from bacteria grown in different physiological conditions (hypoxia, stationary phase, and low temperature) and from different *in vivo* conditions (intestinal lumen and blood), at least two of the sRNAs were found to be involved in virulence. One of these was *rli38*, which was upregulated 25-fold in blood and can pair to three mRNAs including that for *fur*, which encodes a global iron-uptake regulator. A subsequently constructed *rli38* mutant was attenuated in the murine listeriosis infection model. Significantly, *rli38* is absent from *Listeria innocua,* a nonpathogenic species closely related to *L. monocytogenes.* To assess transcriptional reshaping further, Toledo-Arana *et al.* [6] also investigated the transcriptomes of bacteria mutant for the global regulators PrfA (a virulence determinant regulator), SigB (an alternative sigma factor) and Hfq (an RNA chaperone). Together, these approaches identified the differential regulation of several additional RNA elements including *cis*-regulatory RNAs and overlapping UTRs, providing a comprehensive picture of how the transcriptome changes as the microorganism cycles through saprophytism and virulence. The results of the *in vivo* transcriptional profiling were largely verified in a study by Camejo *et al.* [7] using a *L. monocytogenes* macroarray.

Whereas the microarray studies described above give unique insights in the biology of *L. monocytogenes* (as did previous tiling microarray studies on *E. coli* [22] and *Caulobacter crescentus* [23]), we may have now started to reach the limits of what microarray and other hybridization-based techniques can tell us [24]. Microarrays have a relatively limited dynamic range for the detection of transcript levels, owing to background, cross-hybridization, saturation and spot density and quality, and require a complete genome sequence for probe design. Also, comparison of transcription levels between independent microarray experiments is rather challenging and mostly based on complex normalization methods [25]. Finally, microarray technology only measures the relative level of RNA, and does not distinguish between *de novo* synthesized transcripts and modified transcripts; nor does it allow accurate determination of the promoter used in the case of *de novo* transcription. Many of these issues can now be resolved by high-throughput sequencing of cDNA libraries [24].

This is illustrated in recent work by Yoder-Himes *et al.* [18], who used high-throughput sequencing to compare transcriptional patterns between two *Burkholderia cenocepacia* strains in conditions mimicking cystic fibrosis (CF) and soil, using strains originally isolated from a patient with CF and from soil. They also report significant changes in the transcription of noncoding RNA that are important in adaptation to external conditions. Other researchers have described transcriptomics or noncoding RNA identification using high-throughput sequencing in the bacterial pathogens *Bacillus anthracis* and *V. cholerae* [19,20], while alternative approaches, such as immunoprecipitation of nucleic-acid-binding proteins, have been used to identify RNAs bound to the RNA chaperone Hfq in *Salmonella* [21].

Many regulatory noncoding RNAs were first described in *E. coli* and related Enterobacteriaceae, and depend on Hfq for their mode of action. However, with the advent of high-density microarrays and high-throughput sequencing, it is clear that paradigms based on *E. coli* do not necessarily hold true in other bacteria. Some bacteria lack Hfq and can be predicted to use alternative mechanisms for RNA regulation, while in many Hfq-positive bacteria the role of Hfq is still unclear [16]. Other RNA-binding proteins, such as CsrA, may play important roles beyond their originally described functions; in the case of CsrA this is binding to Shine-Dalgarno-like sequences [9]. All these regulatory pathways taken together allow for fine-tuning of the cellular transcriptional mechanisms, giving the bacterium the best of options to survive in adverse conditions.

It is now clear that, as in eukaryotes, bacteria exploit noncoding RNAs in their genetic regulatory repertoire, destroying another myth in the distinction between eukaryotes and prokaryotes. The development of 'next generation' DNA sequencing and direct RNA sequencing will no doubt throw even more light on the role of RNA in gene regulation [26,27]. This will pave the way to a new era in understanding the complex and dynamic mechanisms by which bacteria adapt to the multiple environments they encounter. The next few years promise to be a voyage of discovery in terms of understanding the previously underestimated role of RNA in bacterial gene regulation.

Acknowledgements

References
1. Cossart P, Toledo-Arana A: ***Listeria monocytogenes*, a unique model in infection biology: an overview.** *Microbes Infect* 2008, **10:**1041-1050.

2. Bijlsma JJ, Groisman EA: **Making informed decisions: regulatory interactions between two-component systems.** *Trends Microbiol* 2003, **11:**359-366.

3. Waters LS, Storz G: **Regulatory RNAs in bacteria.** *Cell* 2009, **136:**615-628.

4. Winkler WC, Breaker RR: **Regulation of bacterial gene expression by riboswitches.** *Annu Rev Microbiol* 2005, **59:** 487-517.

5. Vogel J: **A rough guide to the noncoding RNA world of *Salmonella*.** *Mol Microbiol* 2008, **71:**1-11.

6. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, Loh E, Gripenland J, Tiensuu T, Vaitkevicius K, Barthelemy M, Vergassola M, Nahori MA, Soubigou G, Régnault B, Coppée JY, Lecuit M, Johansson J, Cossart P: **The *Listeria* transcriptional landscape from saprophytism to virulence.** *Nature* 2009, **459:**950-956.

7. Camejo A, Buchrieser C, Couve E, Carvalho F, Reis O, Ferreira P, Sousa S, Cossart P, Cabanes D: ***In vivo* transcriptional profiling of *Listeria monocytogenes* and mutagenesis identify new virulence factors involved in infection.** *PLoS Pathog* 2009, **5:**e1000449.

8. Wassarman KM: **6S RNA: a regulator of transcription.** *Mol Microbiol* 2007, **65:**1425-1431.

9. Lucchetti-Miganeh C, Burrowes E, Baysse C, Ermel G: **The post-transcriptional regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages of infection in animal hosts.** *Microbiology* 2008, **154:**16-29.

10. Narberhaus F, Waldminghaus T, Chowdhury S: **RNA thermometers.** *FEMS Microbiol Rev* 2006, **30:**3-16.

11. Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, Breaker RR: **Riboswitches in eubacteria sense the second messenger cyclic di-GMP.** *Science* 2008, **321:**411-413.

12. Yanofsky C: **The different roles of tryptophan transfer RNA in regulating *trp* operon expression in *E. coli* versus *B. subtilis*.** *Trends Genet* 2004, **20:**367-374.

13. Van Melderen L, Saavedra De Bast M: **Bacterial toxin-antitoxin systems: more than selfish entities?** *PLoS Genet* 2009, **5:**e1000437.

14. MacLellan SR, Smallbone LA, Sibley CD, Finan TM: **The expression of a novel antisense gene mediates incompatibility within the large repABC family of alpha-proteobacterial plasmids.** *Mol Microbiol* 2005, **55:**611-623.

15. Mattick JS: **The genetic signatures of noncoding RNAs.** *PLoS Genet* 2009, **5:**e1000459.

16. Valentin-Hansen P, Eriksen M, Udesen C: **The bacterial Sm-like protein Hfq: a key player in RNA transactions.** *Mol Microbiol* 2004, **51:**1525-1533.

17. Cloonan N, Grimmond SM: **Transcriptome content and dynamics at single-nucleotide resolution.** *Genome Biol* 2008, **9:**234.

18. Yoder-Himes DR, Chain PS, Zhu Y, Wurtzel O, Rubin EM, Tiedje JM, Sorek R: **Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing.** *Proc Natl Acad Sci USA* 2009, **106:**3976-3981.

19. 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.

20. Passalacqua KD, Varadarajan A, Ondov BD, Okou DT, Zwick ME, Bergman NH: **Structure and complexity of a bacterial transcriptome.** *J Bacteriol* 2009, **191:**3203-3211.

21. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JC, Vogel J: **Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq.** *PLoS Genet* 2008, **4:**e1000163.

22. Selinger DW, Cheung KJ, Mei R, Johansson EM, Richmond CS, Blattner FR, Lockhart DJ, Church GM: **RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array.** *Nat Biotechnol* 2000, **18:**1262-1268.

23. McGrath PT, Lee H, Zhang L, Iniesta AA, Hottes AK, Tan MH, Hillson NJ, Hu P, Shapiro L, McAdams HH: **High-throughput identification of transcription start sites, conserved promoter motifs and predicted regulons.** *Nat Biotechnol* 2007, **25:**584-592.

24. Hoen PAT, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX, Boer JM, van Ommen GJ, den Dunnen JT: **Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms.** *Nucleic Acids Res* 2008, **36:**e141.

25. Hinton JC, Hautefort I, Eriksson S, Thompson A, Rhen M: **Benefits and pitfalls of using microarrays to monitor bacterial gene expression during infection.** *Curr Opin Microbiol* 2004, **7:**277-282.

26. MacLean D, Jones JDG, Studholme DJ: **Application of 'next-generation' sequencing technologies to microbial genetics.** *Nat Rev Microbiol* 2009, **7:**287-296.

27. Morris DR: **Ribosomal footprints on a transcriptome landscape.** *Genome Biol* 2009, **10:**215.