Review

Genetics and Molecular Biology

Small regulatory RNA molecules in bacteria

NG Ternan*

Abstract

Introduction

Small non-coding RNA molecules are widespread in all kingdoms of life, where they serve to regulate and fine-tune gene expression. They can act in *cis* or *trans*, depending upon their structural relationship with genes whose expression they influence and function by interacting with target messenger RNA molecules to inhibit or accelerate translation. Thus, they can exert rapid control on cellular protein levels. Within bacteria, many small RNAs have been described in Gram-negative model organisms, but developments in our understanding of their role in Gram-positive organisms have been slower. It is clear that small RNAs (sRNAs) influence a wide range of cellular processes, including adaptation to environmental stresses, and virulence processes in pathogens. The aim of this review was to discuss the key elements of sRNA biology and to summarise what is known of their role in Clostridia.

Discussion

Historically, identification of small RNAs has been challenging but recent developments in sequencing technology and computational analysis have led to over 45,000 predicted small RNAs being catalogued in the past few years. However, many of these *in silico* predictions are yet to be validated and the complexity, in terms of small RNA interactions with gene networks, means we are really only beginning to understand how wide-ranging their effects can be within bacteria. It is clear that small RNAs play a critical role in all aspects of bacterial physiology. Within the genus *Clostridium*, the role of small RNAs in the pathogens *Clostridium perfringens*, *Clostridium botulinum* and *Clostridium difficile* is much less well understood, despite hundreds of small RNAs having been predicted within these organisms.

Conclusion

RNOmics is a rapidly expanding field and it is clear that advances in our understanding must deploy high-throughput post-genomic technologies such RNA sequencing in efforts to determine the functions of individual bacterial sRNAs. Using in silico predictions as a platform for novel discoveries, it will be of interest to determine the conditions under which sRNAs are expressed, and whether strain to strain variations exist. The research community now has an opportunity to identify, and consequently to define the roles of, sRNAs within these microorganisms.

Introduction

The genus *Clostridium* encompasses a heterogeneous group of Gram-positive endospore-forming obligately anaerobic microorganisms that are ubiquitous in soils and the intestines of higher organisms. Certain species are economically useful and relatively benign, for example, *Clostridium acetobutylicum*, *Clostridium beijerinckii* and *Clostridium cellulovorans* are employed in the industrial production of biofuels[1]. However, the genus is also infamous for the toxin-producing pathogens *C. difficile*, *C. perfringens* and *C. botulinum*, whose yearly socioeconomic impact is considerable[2,3]. *Clostridium difficile* infection (CDI) causes infectious diarrhoea with associated abdominal pain, cramping and low grade fever up to 40.6°C[4]. *C. difficile* pathogenesis and many of the factors underlying CDI are well understood, but CDI can still be life-threatening if not treated promptly[2,4,5].

The availability of over 30 *C. difficile* genome sequences[6–9] has afforded researchers excellent opportunities to better understand the evolution and lineages of these organisms. Generation of comparative functional genomic data sets has lagged somewhat, and as a consequence, comparatively little is known about the adaptive ability of *C. difficile*. Thus, in our laboratory, we have taken a systems biology approach to understanding the response of *C. difficile* to clinically relevant heat stress, using comparative proteomics and transcriptomics[10,11]. While a classical heat-shock response and class I chaperone induction was observed at 41°C, we also observed downregulation of the flagellum, *FliC* (CD0239) and several other recognised virulence factors, such as *cwp20* (CD1469), *cwp5* (CD2786) and *TcdA* (CD0663), strengthening the hypothesis that virulence of *C. difficile* is 'set' at 37°C. We also determined that the correlation between changes in protein abundances and their cognate transcripts was inconsistent. Several factors could explain this observation including protein/mRNA stability, transcription efficiency or unrecognised post-transcriptional regulatory mechanisms[12]. Recently, Chen et al.[13] demonstrated the presence of small, non-coding, regulatory RNA molecules (small RNAs, sRNAs) in *C. acetobutylicum* and proposed a role for them in gene regulation in this microorganism. This review seeks to provide an overview of the key elements of sRNA biology and to summarise what is known of their role in Clostridia.

* Corresponding author
Email: ng.ternan@ulster.ac.uk

Northern Ireland Centre for Food and Health, School of Biomedical Sciences, University of Ulster, Coleraine, Northern Ireland

Review

Discussion

The author has referenced some of his own studies in this review. The protocols of these studies have been approved by the relevant ethics committees related to the institution in which they were performed.

What are small RNAs?

Eukaryotic sRNAs

Small non-coding RNAs, including microRNAs (miRNAs) and small interfering RNAs (siRNAs), have been identified as regulators of a variety of cellular processes in plants and animals[14]. First described in *Caenorhabditis elegans*[15] several hundred miRNAs, generally ~21–22 nucleotides in length, have now been described. They are generated by cleavage of longer, precursor RNA transcripts that have formed a self-complementary foldback loop by the RNAseIII-like enzyme Dicer and function by base pairing with target mRNA, initiating its degradation. siRNAs (<30 nucleotides) are generated by Dicer-mediated cleavage of double-stranded RNA (dsRNA) and play a role in RNA interference (RNAi) via the RNA-induced silencing complex, where they guide sequence-specific cleavage of RNAs. Thus, these sRNAs are functionally interchangeable. miRNAs have also been predicted and experimentally verified in DNA viruses, with the herpesviridae containing the largest number of viral miRNAs: for such viruses that undergo persistent infection, the invisibility of miRNAs to the adaptive immune response is a useful trait[16,17]. Indeed, the fact that dsRNAs are quite stable *in vivo* and non-immunogenic means that RNAi has great potential for therapeutic use[18,19]. It is known that miRNAs encoded by both host and infecting viruses enable these protagonists to battle with each other during infection[20], and as a result, miRNA profiles are becoming recognised as novel means of diagnosis[21].

Prokaryotic sRNAs

SRNAs were initially identified in bacteria with the identification of 6s RNA in *Escherichia coli*[22], but it is only relatively recently that their influence on bacterial cellular processes and their varied modes of action have become recognised[23]. In contrast to shorter eukaryotic or viral miRNAs, bacterial sRNAs (i.e. not tRNA, rRNA or 5S RNA) are typically between 50 and 500 nucleotides in length, and as with many developments in microbial sciences, *E. coli* was, and is, the model organism for study of sRNAs. Knowledge of sRNA biology in Gram-positive organisms and in archaea has developed more slowly, due in part to a lack of efficient genetic tools[24,25]. Initial elucidation of an individual sRNA's function in *E. coli* came in 1984. Mizuno and colleagues[26] showed that an mRNA-interfering complementary RNA that was complementary to the 5'-end region of the *ompF* gene mRNA, served to inhibit production of the ompF protein by interfering with translation. This new field of RNA biology has been dubbed RNOmics[27]. It has since developed and expanded exponentially, assisted in no small part by technical advances in DNA sequencing technologies and the development of computational algorithms for identification of sRNA sequences in genomic information (Figure 1).

Functionality of bacterial sRNAs

Bacterial sRNAs regulate and fine-tune gene expression in bacteria and it is thought that they enable a faster response to changing conditions at relatively low metabolic cost. Functional RNA molecules require only limited transcription energy compared with other cellular regulatory mechanisms, and in addition, less time is required for an sRNA to be produced and to impact upon target protein levels[28]. A wide range of environmental stimuli impact upon sRNA expression and it is not surprising that many sRNAs are associated with bacterial stress responses[29]. sRNAs can exert global effects on gene expression. In the oxidative stress response in *E. coli*, for example, the 109-nucleotide OxyS sRNA is transcribed divergently from, and regulated by, the *oxyR* gene encoding the redox-sensitive transcriptional regulator that is the actual sensor of the oxidative shock. Upon expression of oxyS sRNA, translation of rpoS is inhibited with rapid and global
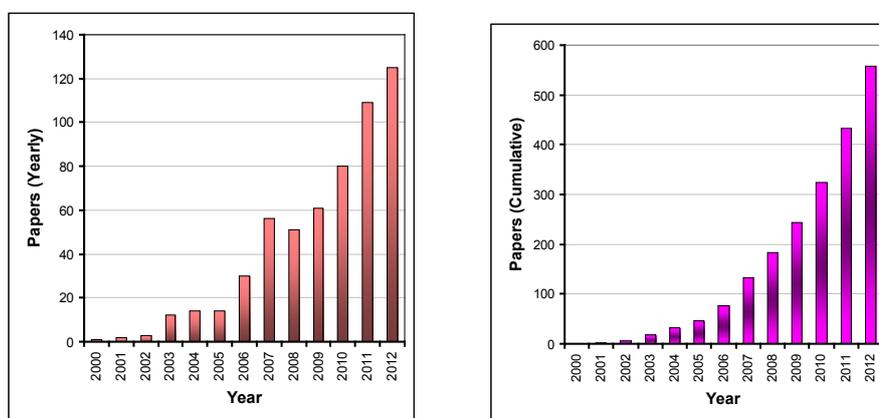

Figure 1: The rise in RNomics over the last decade. Year on year (a) and cumulative total (b) publications from the Web of Knowledge database that contain both 'sRNA*' and 'bacteria*' in the title. It is clear that we are in exponential phase of this new and exciting research area and also that there is a significant body of literature.

effects upon cellular physiology[29,30]. In *E. coli*, *FlhDC*–the master regulator of flagellar biosynthesis–is regulated by multiple protein transcription factors that respond to different environmental stimuli including cell envelope stress and salt concentration. However, the recent work of De Lay and Gottesman[31] has shown that complexity, and thus regulatory power, is increased because the 5' untranslated region (5' UTR) of the *flhDC* mRNA is also subject to negative regulation by six different sRNA molecules (ArcZ, OmrA, OmrB, OxyS, SdsR and GadY) and positive regulation by one (McaS). Thus, the *flhDC* mRNA serves as a hub that allows integration of signals derived from environmental salt and oxygen concentrations, oxidative insult and the general stress response into the decision to make flagella. The question of whether the flagellum is a primary *C. difficile* virulence factor is open to debate[11], but a flagellar filament requires some 2% of a bacterial cell's total energy consumption under optimal growth conditions, in order to synthesise the necessary ~20,000 subunits of FliC protein: it is clear why such precise regulation of flagellar biosynthesis might be necessary. It has been suggested that up to ~300 sRNAs will be present in the average bacterial genome, a number equivalent to the complement of transcription factors[32]. As exemplified above, however, these sRNAs have many times the potential regulatory capacity of protein transcription factors, and thus they are clearly of critical importance in bacterial physiology.

How do small RNA molecules exert their biological effect?

In the Gram-positive bacterial pathogens in which sRNAs have been characterised to date, their biological functions have been linked to adaptation or virulence. For example, in *C. perfringens*, the VR-RNA sRNA regulates collagenase and alpha toxin gene transcription[33]. Like Gram-negatives,

Gram-positive bacteria have many sRNA-mediated regulatory mechanisms that allow response to environmental and intercellular signals via a number of different mechanisms[24]. Bacterial sRNAs are generally found in the intergenic regions of the genome, and they fall into two main categories depending upon their genomic context in relation to the target gene. Those that are transcribed independently from the target gene are encoded in *trans*, while those that are co-transcribed, usually from within the 5' UTR of the target transcript, are encoded in *cis*[29] (Figure 2). *Cis*-encoded sRNAs can also be transcribed from the antisense strand at the same genetic locus as the target, and these antisense RNAs will therefore exhibit perfect complementarity with their target, allowing interactions that impact positively or negatively upon gene expression[32,34]. *Cis* and *trans* sRNAs can be further categorised into two subgroups based on their mode of action. Certain sRNAs pair with mRNA targets to affect their stability or translation, whereas others act as molecular decoys that bind to protein targets and affect their activity[35–39] (Figure 3). RNA thermosensors (Figure 4a) have been demonstrated to play pivotal regulatory roles in not only the heat stress response but also in the co-

ordination of expression of virulence genes in number of human pathogens[40,41] while Riboswitches, a further class of *cis*-acting RNA element, control expression of downstream genes via metabolite-induced alteration of sRNA secondary structures (Figure 4b). Riboswitches can function in a variety of ways but in brief, different metabolites can allow them either to induce or repress transcription or translation, as recently reviewed by Serganov and Nudler[42]. The bacterial sRNAs that have been characterised in Gram-positive microorganisms are expressed mainly in a growth phase-dependent manner, and while it may be hypothesised that, like in *E. coli*, they are part of complex regulatory processes our current knowledge of factors affecting sRNA expression in Gram-positive bacteria is lacking[24]. Thus, while sRNAs have been characterised in *Bacillus subtilis*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Streptococcus pyogenes*, *C. acetobutylicum* and *C. perfringens*, very little is known about their role in *C. difficile*.

The role of the Hfq RNA chaperone protein

While sRNA modes of action are fairly similar between Gram-positive and Gram-negative bacteria, one aspect of sRNA biology that is
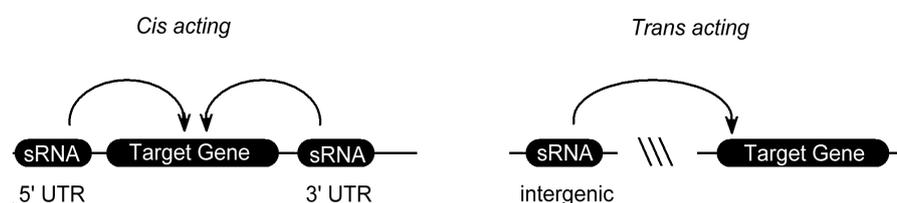
Figure 2: Generalised genomic context of *cis*- and *trans*-acting small RNAs. Cis-acting sRNAs are generally found in the 5'UTR of the mRNA (5'UTR), although less commonly, they may be encoded in the 3'UTR. Riboswitches and RNA thermometers fall into this class of sRNAs. Trans-acting sRNAs are encoded in intergenic regions of the genome (characterised by the presence of rho-independent terminators and promoters in their sequence) and are transcribed independently of the target. They usually act by base pairing (often assisted by the Hfq RNA chaperone protein) with the target mRNA, influencing the output from that mRNA.
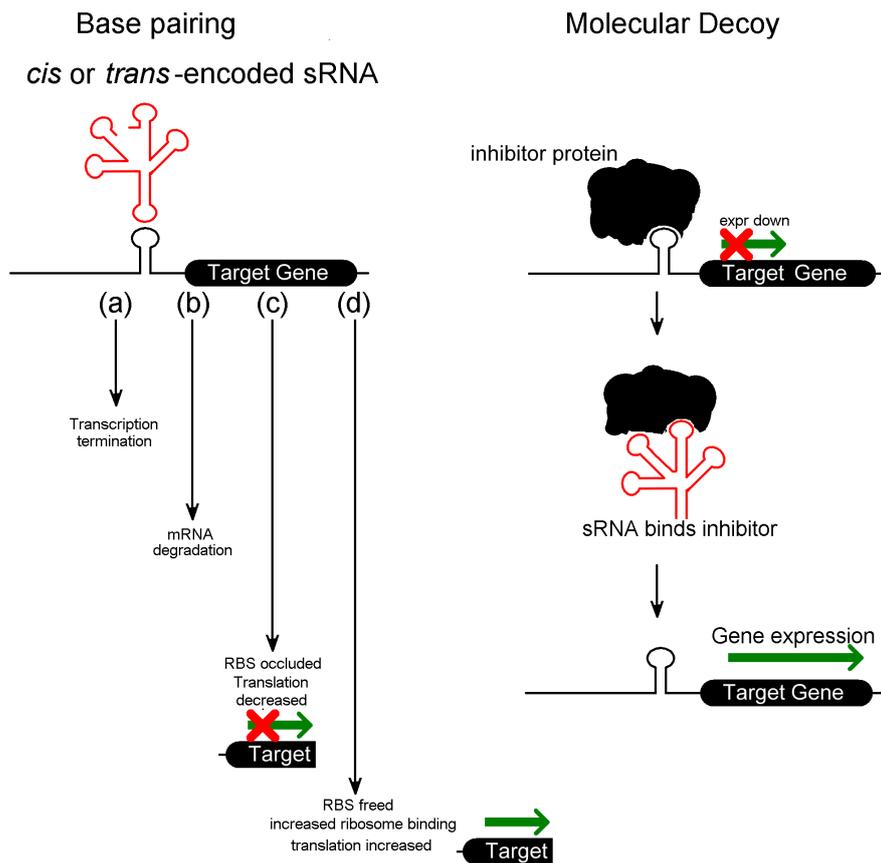
Review

Figure 3: Small RNA molecules can act to modulate gene expression in a variety of ways. Base pairing of the sRNA with a target mRNA sequence can lead to (a) termination of transcription, (b) degradation of the mRNA, (c) occlusion of the ribosome binding site (RBS) and decreased translation or (d) changes in the secondary structure of mRNA such that the RBS is more accessible by the 30S ribosome and translation is increased. In an alternative mechanism, the (*trans*-encoded) sRNA acts as a molecular decoy–here, binding of an inhibitor protein to the mRNA prevents translation but if the inhibitor is sequestered by binding to the decoy sRNA, repression is lifted.

less well conserved is the role of the Hfq RNA chaperone. Hfq is highly conserved in prokaryotes and belongs to the Sm family of proteins that are known to interact with RNA in both eukaryotes and prokaryotes[43]. Hfq has been shown to interact with a considerable number of *trans*-encoded sRNA molecules in Gram-negative microbes, where it plays a key role in stabilising sRNA molecules or facilitating interaction with mRNA targets[24,44]. Thus, Hfq plays a key role in one of the most complex post-transcriptional networks known[45]. In low GC Gram-positive bacteria, however, the function of Hfq is still unclear[43], although in *L. monocytogenes*, Hfq is required for function of several sRNAs (LhrA–C)[46]. However, other *L. monocytogenes* sRNAs do not require Hfq for target interaction[47], and in *S. aureus*, Hfq does not seem to be required for sRNA–mRNA interactions at all[48]. There is also the consideration that not all bacterial genomes contain an Hfq homologue, raising the possibility that other proteins may be able to substitute for Hfq in certain organisms[49].

Identification and validation of sRNAs in bacteria
Initial identification of sRNAs in bacteria is challenging, not least because until recently there was no general approach that provided a comprehensive solution to their prediction[21]. Furthermore, sRNA target prediction is awkward because many sRNA:mRNA hybridisations occur over relatively short regions of imperfect complimentarity[50]. The initial work on sRNAs some 40 years ago used gel electrophoresis to fractionate radiolabelled total bacterial RNA, followed by elution of low-molecular-mass RNA molecules from the gels and subsequent analysis[51]. In the 30 years since their first discovery, only around a dozen sRNAs were identified and characterised in *E. coli*, but since then, developments in genomics and computational biology have allowed the field of sRNA biology to expand massively. In the past decade or so, sRNA gene finders based on well-characterised sequences and algorithms to predict the minimum free energy of structured RNAs have been applied to newly catalogued bacterial genomes[52,53]. In addition, comparative genomic approaches that allow researchers to make sRNA predictions based on the presence of rho-independent terminators and promoters and other features in the intergenic regions have also been used to predict sRNAs[13,50,54].

A workflow for sRNA characterisation, therefore, might proceed from *in silico* identification of sRNAs to demonstration of their expression by qRT-PCR or Northern blotting and the subsequent identification of direct and indirect targets of individual sRNA molecules using *in silico* prediction algorithms followed by wet laboratory methods to validate the interactions. For example, in the work of Chen et al.[13], the only
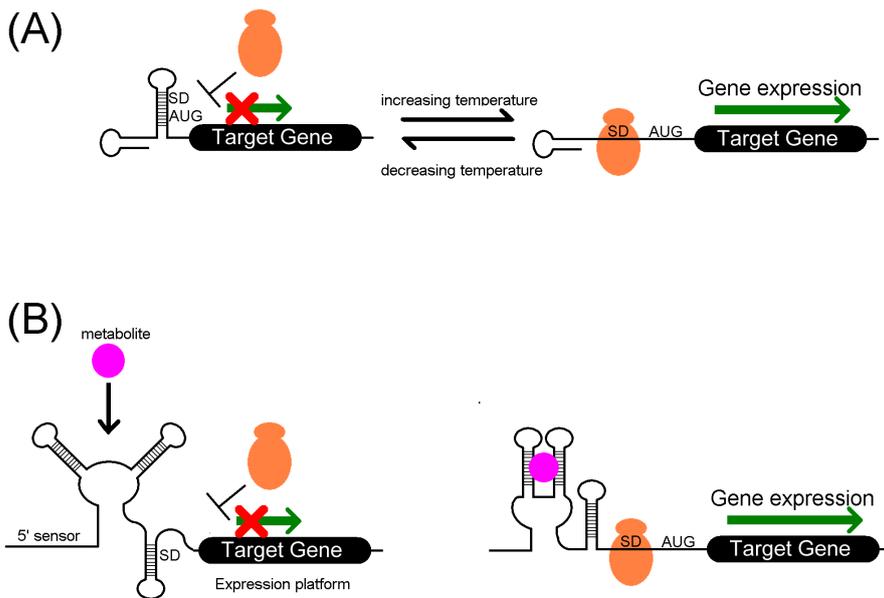
Figure 4: RNA thermometers and Riboswitches are examples of *cis*-encoded small RNA molecules. (a) At low temperature, the 30S ribosome is prevented from accessing the shine dalgarno (SD) sequence and the start codon (AUG) due to the complex secondary structure of the mRNA. Upon increasing temperature, the secondary structure gradually melts and the ribosome can access the SD and AUG. This is thus a faster, direct, temperature sensing mechanism which is known to regulate heat-shock gene expression and virulence in bacteria. Sequence conservation in the 5' aptamer domain enables database searches for identification of these thermosensing elements. (b) Generalised mechanism for expressional control via metabolite binding to *cis*-acting riboswitches. The riboswitch consists of a sensor aptamer domain which can bind the metabolite (e.g. anions, metal ions, cofactors, purines and amino acids are all known to direct switching) and an expression platform. Riboswitches sense different concentrations of a single metabolite, and upon highly discriminatory binding of the metabolite to the aptamer domain, the secondary structure of the element changes to allow changes in transcription, translation, splicing and mRNA stability.

post-genomic fields, for example mass spectrometry-driven proteomics[57], these developments have been crucial to the expansion of sRNA biology as a field of research. Within even the past few years, the number of sRNAs identified in a wide range of bacteria, including in Gram-positives, has increased at an incredible rate. A natural consequence of this success is an increasing urgency for identification of their cellular targets and functional roles, a facet of the research that has lagged considerably behind identification studies[58]. A number of groups have presented a variety of tools for the purposes of sRNA identification[59]. One of the longest standing is the Rfam database[60], a collection of non-coding RNA families represented by multiple-sequence alignments and secondary structure predictions that was first developed a decade ago[60,61]. The work of Livny et al.[50] introduced the powerful *s*RNA *i*dentification *p*rotocol using *h*igh-throughput *tech*nologies (SIPHT) tool, which incorporates a number of programmes and adjustable search parameters to identify sRNAs and other features in an automated fashion. SIPHT identifies conserved sequences along with rho-independent terminators and promoters in intergenic regions and incorporates BLAST, genomic synteny and transcription factor-binding site analyses into a workflow that yields an output that can be opened in Excel. This work has allowed prediction of candidate sRNA-encoding loci from over 900 bacterial genomes and plasmids within the NCBI database, thus expanding the number of predictions from several hundred candidate sRNAs to over 45,000. However, all databases will have perceived drawbacks, regardless of how they are implemented. They might not allow further analysis, or they may be restricted to a limited number of bacterial species, or be reliant on published data. Two recent publications have sought to redress

report to date on genome-wide characterisation of sRNAs in clostridia, *in silico* methods were used to predict sRNAs in 21 clostridial species. The authors then used qRT-PCR to validate 30 sRNAs of 113 predicted in *C. acetobutylicum* and 21 from *C. botulinum*, thus showing that qRT-PCR is a useful first screening step. Highly expressed sRNAs (by qRT-PCR) were then analysed using Northern blotting to validate transcript sizes against those predicted by the *in silico* analysis. A number of additional experimental approaches can also be used including tiling oligonucleotide microarrays, cDNA cloning and high-throughput RNAseq[54–56]. In addition, the identification of sRNA:Hfq associations can provide further evidence that transcripts are sRNAs[45].

Databases for sRNA research
Concomitantly with these predictive methods and experimental validations, the development of user-friendly, browser-based databases and software tools to allow information retrieval and analysis has proceeded apace. As with other

this deficiency: sRNAdb, developed by Pischimarov and colleagues[59], is a user-friendly searchable database allowing comprehensive comparative analysis of sRNAs from Gram-positive microorganisms. In addition, the end user may incorporate further features of interest into a local customised database. The work of Li et al. describes BSRD–a repository for bacterial small regulatory RNA[62], which is said to contain more experimentally validated sRNAs than any other database and enables researchers to identify and characterise sRNAs in large-scale transcriptome sequencing projects. Thus, researchers interested in a particular bacterial group now have at their disposal a comprehensive range of predictions, databases and *in silico* analysis tools to underpin their investigations.

Identification of small RNA targets– dissection of roles and functions

Having validated the existence of a population of sRNAs, there remains the issue of what individual sRNA molecules actually do. It is clear that only a relatively small proportion of the sRNAs predicted to date have had their targets experimentally verified, although targets can initially be inferred computationally. Many sRNAs are antisense regulators, and bioinformatics searches for complementarity can assist with target identification–although in reality, the base pairing between sRNAs and their targets is often imperfect, making this task difficult[58]. One such tool, sTarPicker, is based on a mathematical model of hybridisation between sRNA and mRNA and is said to predict sRNA targets with higher efficiency than competing programmes[63]. sRNATarBase, developed by the same group, seeks to provide a resource of sRNA targets that have been experimentally verified, thus providing support for predictive models and subsequent *in silico* and functional analyses. The authors systematically and manually

collected sRNA:target interaction data from published papers in order to develop their database of sRNA targets[64]. However, where targets are as yet only inferred, it is still necessary to validate these sRNA:target predictions and to this end, several interesting approaches can be used. In addition, the determination of what constitutes a primary target (direct interaction with the sRNA) and what is a secondary target, such as a transcription factor, is also of considerable importance[58].

Analysis of the sRNA and proposed target mRNA expression under different conditions is one approach to target identification. As reported by Chen et al.[13], a conserved novel sRNA (CAC610) in *C. acetobutylicum* and a downstream gene (CAC0528) both responded to the antibiotic clindamycin. As the distance between the sRNA and the gene was conserved across a number of clostridial strains at ~185 bp (although neither exist within in *C. difficile*), the authors concluded that there was a functional relationship between the two, although the exact mechanism by which the sRNA might modulate gene expression (or *vice versa*) was not determined[13]. Another method for determination of sRNA targets has been described as a 'biochemical fishing expedition'. The use of sRNA molecules as the bait in order to capture an mRNA target is an approach that can be further refined by incorporating a recombinant affinity tagged Hfq protein. As many sRNAs interact with Hfq, its subsequent purification, complete with sRNA and the sRNA target, can allow sRNA target identification. In this instance, creation of cDNA clones, and their hybridisation to whole-genome microarrays, could be employed[65]. Functional genomic analyses, for instance, with mutants constructed in validated sRNA-encoding regions of the genome, allow the subsequent determination of the effect of these deletions on both host cell physiol-

ogy and on the expression of predicted targets[66,67]. With mutants in hand, tiling oligonucleotide microarrays, or RNAseq analysis, would provide a genome-wide picture of their effect. Furthermore, it should be possible to experimentally express a high level of a given sRNA in a host cell and compare global cellular responses with those of either the wild-type or a deletion mutant[68].

Conclusion

RNOmics is still a rapidly expanding field, and it is clear that advances in our understanding must be driven by the use of high-throughput post-genomic technologies such as transcriptome sequencing. Efforts will also be required to determine the functions of individual bacterial sRNAs, a not inconsiderable task, given the potential for widespread interactions of sRNAs with multiple targets and within gene networks. ithin the clostridia, there is still much to be done to experimentally validate sRNA predictions and interestingly, it appears that the number of sRNAs is related to the physiology of the organism. Greater numbers of sRNAs have been predicted in the genomes of pathogenic clostridia–for example, *C. difficile* 630 is predicted to contain 264 sRNAs, few of which have been experimentally verified as yet. Clostridial sRNAs appear to be phylogenetically restricted to these organisms and are not conserved in, for example, Bacilli; thus, it will be of interest to determine precisely under what conditions these sRNAs are expressed, and whether strain to strain variations exist. An obvious question, from an epidemiological perspective, would be whether sRNAs' expression varies both within and–between different *C. difficile* ribotypes, for example.

The work so far on *C. acetobutylicum* suggests that certain sRNAs may play a role in antibiotic resistance, and this observation provides new avenues for research into

antibiotic tolerance mechanisms, drug targets and diagnostic methods. At present, there are no data on the role of the Hfq homologues that exist in the genomes of *Clostridium* spp., although with modern functional genomics tools it should be possible to construct gene knockouts and determine the role of Hfq. Our understanding of sRNAs in Clostridia is at present incomplete, presenting the research community with an opportunity to identify, and consequently, to define the roles of these RNAs. within these anaerobic microorganisms.

Abbreviations list

5' UTR, 5' untranslated region; CDI, *Clostridium difficile* infection; dsRNA, double-stranded RNA; miRNAs, microRNAs; RNAi, RNA interference; sRNAs, small RNAs; siRNAs, small interfering RNAs.

References

1. Ezeji TC, Qureshi N, Blasckek HP. Industrially relevant fermentations. In: Dürre P, editor. Handbook on Clostridia. Boca Raton, Florida: CRC Press, Boca Raton, Florida; 2005. pp.797–812.

2. Rupnik M, Wilcox MH, Gerding DN. Clostridium difficile infection: new developments in epidemiology and pathogenesis. Nature Rev Microbiol. 2009 July;7(7): 526–36.

3. Peck MW, Stringer SC, Carter AT. Clostridium botulinum in the post-genomic era. Food Microbiol. 2011 Apr;28(2): 183–91.

4. Bartlett JG. Historical perspectives on studies of Clostridium difficile and C. difficile infection. Clin Infect Dis. 2008 Jan;46(Suppl 1):S4–11.

5. Pépin J, Saheb N, Coulombe MA, Alary ME, Corriveau MP, Authier S, et al. Emergence of fluoroquinolones as the predominant risk factor for Clostridium difficile-associated diarrhea: a cohort study during an epidemic in Quebec. Clin Infect Dis. 2005 Nov;41(9):1254–60.

6. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, et al. The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. Nat Genet. 2006 Jul;38(7):779–86.

7. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009 Sep;10(9):R102.

8. He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, et al. Evolutionary dynamics of Clostridium difficile over short and long time scales. Proc Natl Acad Sci U S A. 2010 Apr;107(16):7527–32.

9. Forgetta V, Oughton MT, Marquis P, Brukner I, Blanchette R, Haub K, et al. Fourteen-genome comparison identifies DNA markers for severe-disease-associated strains of Clostridium difficile. J Clin Microbiol. 2011 Jun;49(6):2230–8.

10. Jain S, Graham C, Graham RL, McMullan G, Ternan, NG. Quantitative proteomic analysis of the heat stress response in Clostridium difficile strain 630. J Proteome Res. 2011 Sep;10(9):3880–90.

11. Ternan NG, Jain S, Sirivastava M, McMullan G. Comparative transcriptional analysis of clinically relevant heat stress response in Clostridium difficile strain 630. PLoS One. 2012 Jul;7(7):e42410.

12. Olivares-Hernández R, Bordel S, Nielsen J. Codon usage variability determines the correlation between proteome and transcriptome fold changes. BMC Syst Biol. 2011 Feb;5:33.

13. Chen Y, Indurthi DC, Jones SW, Papoutsakis ET. Small RNAs in the genus Clostridium. MBio. 2011 Jan–Feb;2(1): e00340–10.

14. Carrington JC, Ambros, V. Role of MicroRNAs in plant and animal development. Science. 2003 Jul;301(5631): 336–8.

15. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998 Feb;391(6669):806–11.

16. Pfeffer S, Zavolan M, Grässer FA, Chien MC, Russo JJ, Ju JY, et al. Identification of virus-encoded microRNAs. Science. 2004 Apr;304(5671):734–6.

17. Kincaid RP, Sullivan CS. Virus-encoded microRNAs: an overview and a look to the future. PLoS Pathog. 2012 Dec;8(12):e1003018.

18. Esquela-Kerscher A, Slack FJ. Oncomirs-microRNAs with a role in cancer. Nat Rev Cancer. 2006 Apr;6(4):259–69.

19. Siomi MC. Short interfering RNA-mediated gene silencing; towards successful

application in human patients. Adv Drug Deliv Rev. 2009 Jul;61(9):668–71.

20. Houzet L, Jeang KT. MicroRNAs and human retroviruses. Biochim Biophys Acta. 2011 Nov–Dec;1809(11–12):686–93.

21. Mraheil MA, Billion A, Kuenne C, Pischimarov J, Kreikemeyer B, Engelmann S, et al. Comparative genome-wide analysis of small RNAs of major Gram-positive pathogens: from identification to application. Microb Biotechnol. 2010 Nov;3(6):658–76.

22. Brownlee GG. Sequence of 6S RNA of E. coli. Nat New Biol. 1971 Feb;229(5): 147–9.

23. Wassarman KM, Zhang AX, Storz G. Small RNAs in Escherichia coli. Trends Microbiol. 1999 Jan;7(1): 37–45.

24. Romby P, Charpentier, E. An overview of RNAs with regulatory functions in gram-positive bacteria. Cell Mol Life Sci. 2010 Jan;67(2):217–37.

25. Prasse D, Ehlers C, Backofen R, Schmitz RA. Regulatory RNAs in archaea: first target identification in Methanoarchaea. Biochem Soc Trans. 2013 Feb;41(1): 344–9.

26. Mizuno T, Chou MY, Inouye M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proc Natl Acad Sci U S A. 1984 Apr;81(7):1966–70.

27. Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jäger JG, Hüttenhofer A, et al. RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria. Nucleic Acids Res. 2003 Nov;31(22):6435–43.

28. Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiol Rev. 2010 Sep;34(5):866–82.

29. Repoila F, Majdalani N, Gottesman S. Small non-coding RNAs, co-ordinators of adaptation processes in Escherichia coli: the RpoS paradigm. Mol Microbiol. 2003 May;48(4):855–61.

30. Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. Cell. 1997 Jul;90(1):43–53.

31. De Lay N, Gottesman S. A complex network of small non-coding RNAs regulate motility in Escherichia coli. Mol Microbiol. 2012 Nov;86(3):524–38.

32. Irnov I, Sharma CM, Vogel J, Winkler WC. Identification of regulatory RNAs in

Bacillus subtilis. Nucleic Acids Res. 2010 Oct;38(19):6637–51.

33. Shimizu T, Yaguchi H, Ohtani K, Banu S, Hayashi H. Clostridial VirR/VirS regulon involves a regulatory RNA molecule for expression of toxins. Mol Microbiol. 2002 Jan;43(1):257–65.

34. Salim NN, Feig AL. An upstream Hfq binding site in the fhlA mRNA leader region facilitates the OxyS-fhlA interaction. PLoS One. 2010 Sep;5(9):e13028.

35. Babitzke P, Romeo T. CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Curr Opin Microbiol. 2007 Apr;10(2):156–63.

36. Gottesman S. Stealth regulation: biological circuits with small RNA switches. Genes Dev. 2002 Nov;16(22):2829–42.

37. Toledo-Arana A, Repoila F, Cossart P. Small noncoding RNAs controlling pathogenesis. Curr Opin Microbiol. 2007 Apr; 10(2):182–8.

38. Podkaminski, D, Vogel, J. Small RNAs promote mRNA stability to activate the synthesis of virulence factors. Mol Microbiol. 2010 Dec;78(6):1327–31.

39. Desnoyers G, Bouchard MP, Massé E. New insights into small RNA dependent translational regulation in prokaryotes. Trends Genet. 2013 Feb;29(2):92–8.

40. Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. An RNA thermosensor controls expression of virulence genes in Listeria monocytogenes. Cell. 2002 Sep;110(5):551–61.

41. Narberhaus F. Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. RNA Biol. 2010 Jan–Feb;7(1):84–9.

42. Serganov A, Nudler E. A decade of riboswitches. Cell. 2013 Jan;152(1–2):17–24.

43. Nielsen JS, Lei LK, Ebersbach T, Olsen AS, Klitgaard JK, Valentin-Hansen P, et al. Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in Listeria monocytogenes. Nucleic Acids Res. 2010 Jan;38(3):907–19.

44. Sobrero P, Valverde C. The bacterial protein Hfq: much more than a mere RNA-binding factor. Crit Rev Microbiol. 2012 Nov;38(4):276–99.

45. Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. EMBO J. 2012 Oct;31(20):4005–19.

46. Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Søgaard-Andersen L, Kallipolitis BH. Identification of small Hfq-binding RNAs in Listeria monocytogenes. RNA. 2006 Jul;12(7):1383–96.

47. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The Listeria transcriptional landscape from saprophytism to virulence. Nature. 2009 Jun;459(7249):950–6.

48. Bohn C, Rigoulay C, Bouloc P. No detectable effect of RNA-binding protein Hfq absence in Staphylococcus aureus. BMC Microbiol. 2007 Feb;7:10.

49. Pandey SP, Minesinger BK, Kumar J, Walker GC. A highly conserved protein of unknown function in Sinorhizobium meliloti affects sRNA regulation similar to Hfq. Nucleic Acids Res. 2011 Feb;39(11):4691–708.

50. Livny J, Teonadi H, Livny M, Waldor MK. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PLoS One. 2008 Sep;3(9):e3197.

51. Griffin BE. Separation of 32P-labelled ribonucleic acid components. The use of polyethylenimine-cellulose (TLC) as a second dimension in separating oligoribonucleotides of '4.5 S' and 5 S from E. coli. FEBS Lett. 1971 Jun;15(3):165–8.

52. Rivas E, Klein RJ, Jones TA, Eddy SR. Computational identification of noncoding RNAs in E. coli by comparative genomics. Curr Biol. 2001 Sep;11(17):1369–73.

53. Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EGH, Margalit H, et al. Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. Curr Biol. 2001 Jun;11(12):941–50.

54. Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev. 2001 Jul;15(13):1637–51.

55. Pichon C, du Merle L, Caliot ME, Trieu-Cuot P, Le Bouguénec C. An in silico model for identification of small RNAs in whole bacterial genomes: characterization of antisense RNAs in pathogenic Escherichia coli and Streptococcus agalactiae strains. Nucleic Acids Res. 2012 Apr;40(7):2846–61.

56. Qu Y, Bi L, Ji X, Deng Z, Zhang H, Yan Y, et al. Identification by cDNA cloning of abundant sRNAs in a human-avirulent Yersinia pestis strain grown under five different growth conditions.

Future Microbiol. 2012 Apr;7(4): 535–47.

57. Graham C, McMullan G, Graham RL. Proteomics in the microbial sciences. Bioeng Bugs. 2011 Jan–Feb;2(1):17–30.

58. Vogel J, Wagner EG. Target identification of small noncoding RNAs in bacteria. Curr Opin Microbiol. 2007 Jun;10(3): 262–70.

59. Pischimarov J, Kuenne C, Billion A, Hemberger J, Cemič F, Chakraborty T, et al. sRNAdb: a small non-coding RNA database for gram-positive bacteria. BMC Genomics. 2012 Aug;13:384.

60. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy, SR. Rfam: an RNA family database. Nucleic Acids Res. 2003 Jan;31(1):439–41.

61. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, et al. Rfam 11.0: 10 years of RNA families. Nucleic Acids Res. 2013 Jan;41(Database issue):D226–32.

62. Li L, Huang D, Cheung MK, Nong W, Huang Q, Kwan HS. BSRD: a repository for bacterial small regulatory RNA. Nucleic Acids Res. 2013 Jan;41(Database issue):D233–8.

63. Ying X, Cao Y, Wu J, Liu Q, Cha L, Li W. sTarPicker: a method for efficient prediction of bacterial sRNA targets based on a two-step model for hybridization. PLoS One. 2011 Jul;6(7):e22705.

64. Cao Y, Wu J, Liu Q, Zhao Y, Ying X, Cha L, et al. sRNATarBase: a comprehensive database of bacterial sRNA targets verified by experiments. RNA. 2010 Nov;16(11):2051–7.

65. Antal M, Bordeau V, Douchin V, Felden B. A small bacterial RNA regulates a putative ABC transporter. J Biol Chem. 2005 Mar;280(9):7901–8.

66. Barra-Bily L, Fontenelle C, Jan G, Flechard M, Trautwetter A, Pandey SP, et al. Proteomics alterations explain phenoptypic changes in Sinorhizobium meliloti lacking the RNA chaperone Hfq. J Bacteriol. 2010 Mar;192(6):1719–29.

67. Kim JN, Kwon YM. Genetic and phenotypic characterization of the RyhB regulon in Salmonella Typhimurium. Microbiol Res. 2013 Jan;168(1): 41–9.

68. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: mutagenesis in Clostridium refined and streamlined. J Microbiol Methods. 2010 Jan;80(1):49–55.