Accelerating Discovery and Functional Analysis of Small RNAs with New Technologies

Lars Barquist and Jörg Vogel

RNA Biology Group, Institute for Molecular Infection Biology, University of Würzburg, D-97080 Würzburg, Germany; email: lars.barquist@uni-wuerzburg.de, joerg.vogel@uni-wuerzburg.de

Annu. Rev. Genet. 2015. 49:15.1–15.28

The *Annual Review of Genetics* is online at genet.annualreviews.org

This article's doi:
10.1146/annurev-genet-112414-054804

Copyright © 2015 by Annual Reviews.
All rights reserved

Keywords

small RNA, noncoding RNA, Hfq, RNA-seq, high-throughput technology, TraDIS, ribosome profiling, phenotype mapping

Abstract

Over the past decade, bacterial small RNAs (sRNAs) have gone from a biological curiosity to being recognized as a major class of regulatory molecules. High-throughput methods for sampling the transcriptional output of bacterial cells demonstrate that sRNAs are universal features of bacterial transcriptomes, are plentiful, and appear to vary extensively over evolutionary time. With ever more bacteria coming under study, the question becomes how can we accelerate the discovery and functional characterization of sRNAs in diverse organisms. New technologies built on high-throughput sequencing are emerging that can rapidly provide global insight into the numbers and functions of sRNAs in bacteria of interest, providing information that can shape hypotheses and guide research. In this review, we describe recent developments in transcriptomics (RNA-seq) and functional genomics that we expect to help us develop an integrated, systems-level view of sRNA biology in bacteria.

INTRODUCTION

The past decade has witnessed the increasing availability and sophistication of high-throughput technologies, which have led to major changes in our understanding of transcriptional and post-transcriptional regulation in bacteria. Once assumed to be relatively simple, bacterial transcriptomes have been shown to exhibit many of the features associated with transcriptional complexity in eukaryotic systems (64, 171, 175, 198), including a large complement of noncoding RNAs. Most of these noncoding RNAs serve regulatory roles, the majority affecting gene expression through RNA:RNA interactions with other transcripts, although some also act through interactions with regulatory proteins. One abundant class of noncoding RNAs, the *trans*-acting small regulatory RNAs (sRNA), has emerged as a pervasive regulator of diverse cellular processes, with established roles in metabolism (11, 113, 140), quorum sensing (6, 127), biofilm formation (114), iron regulation (130), and virulence (21, 61, 72, 140). The major mode of action of these sRNAs is roughly analogous to that of eukaryotic microRNAs, mediated by intermolecular base-pairing interactions between a conserved seed sequence within the sRNA and an imperfectly complementary target sequence within an mRNA. In some lineages, notably many Gammaproteobacteria, these interactions depend on the chaperone protein Hfq (25, 37, 195), which serves to both facilitate intermolecular interactions and stabilize sRNA transcripts. These interactions can lead to either inhibition or enhancement of protein expression, although inhibition appears to be more common (54). The molecular mechanisms underpinning these changes in gene expression are diverse and can operate at the level of translation initiation by modulating ribosome recruitment, at the level of transcript abundance by modulating transcript degradation, or both (**Figure 1**). Indeed, multiple distinct regulatory mechanisms can be employed simultaneously by the same sRNA (23, 48), with complex regulatory consequences.

sRNAs were discovered serendipitously during the investigation of protein coding sequences (58, 196) (**Figure 2**). For example, the sRNA MicF was initially discovered through a phenotype induced by multicopy plasmids constructed for the purpose of mapping the promoter of *ompC*, encoding an outer membrane protein (119). Similarly, the sRNAs OxyS and DsrA were discovered incidentally as short untranslated transcripts produced in the vicinity of transcriptional regulators (177). The sequencing of a number of enterobacterial genomes in the late 1990s and early 2000s enabled the first genome-wide sRNA screens. These screens used computational methods to predict potential sRNAs (97) on the basis of features such as orphan promoters and terminators (2), covariation of potential base pairs in sequence alignments (160), and sequence conservation coupled with microarray expression data (200), collectively describing several dozen sRNAs in *Escherichia coli* (74). Since that time, technical advances in tiling arrays (181), capillary sequencing (194), pyrosequencing (174), and reversible dye-terminator sequencing (24, 87, 88, 155, 186) have led to steadily increasing numbers of proven and suggested sRNAs, with current estimates putting the number of sRNAs expressed by *E. coli* and *Salmonella enterica* at around 300, many of which are species- or genera-specific. Although sRNA discovery outside of these model systems continues to mature, comparable numbers of sRNAs have been detected in a wide range of bacteria (16, 75, 111, 161, 182), and their relatively narrow conservation on the level of primary sequence appears to be a global trend across bacterial phylogeny (96, 143, 182).

Functional characterization of sRNAs, even in model organisms, has lagged behind discovery. Computational predictions of the mRNA targets of sRNAs using thermodynamic models of intermolecular hybridization have improved by incorporating measures of target sequence accessibility (20, 43) and, recently, interaction conservation (203). However, such predictions still have a generally low accuracy (135), likely because the determinants of sRNA binding are both complex and incompletely understood (9, 136). This issue necessitates the use of experimental approaches in

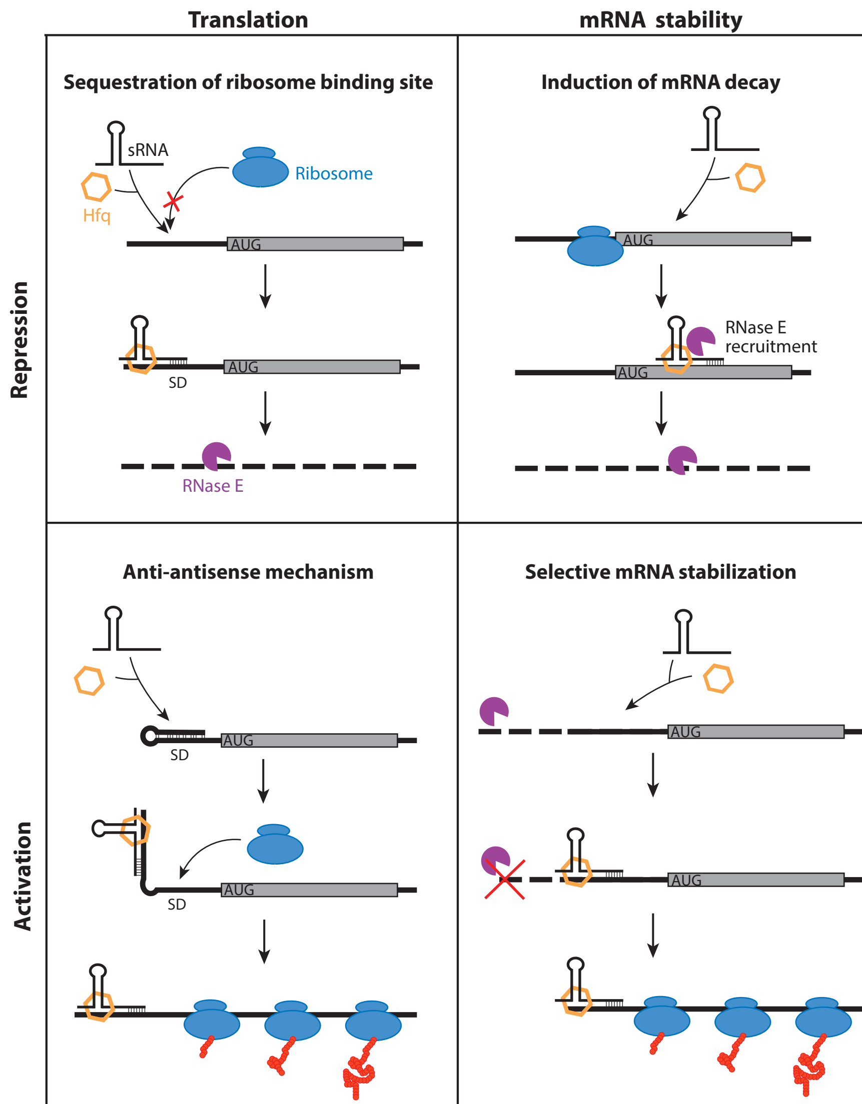

Figure 1

Common mechanisms of sRNA activity. sRNAs can operate by modulating translation (*left*) or transcript stability (*right*), with resulting repression (*top*) or activation (*bottom*) of gene expression. Sequestration of ribosome binding site: sRNA binding in the vicinity of the Shine-Dalgarno (SD) sequence blocks ribosomes, preventing translation of the mRNA. The absence of translating ribosomes protecting the transcript may leave cleavage sites exposed, indirectly leading to transcript degradation. Induction of mRNA decay: An sRNA binding at potentially any position within a target transcript can recruit RNase E, possibly aided by protein:protein interactions between Hfq and RNase E (123). This lowers transcript levels and hence protein production. Anti-antisense mechanism: sRNAs can activate translation by disrupting native mRNA secondary structures that would otherwise block ribosome binding. Selective mRNA stabilization: sRNA binding within a transcript can interfere with degradation, leading to stabilization of particular transcript isoforms (53) or enhanced translation of particular cistrons within a polycistronic transcript (139).

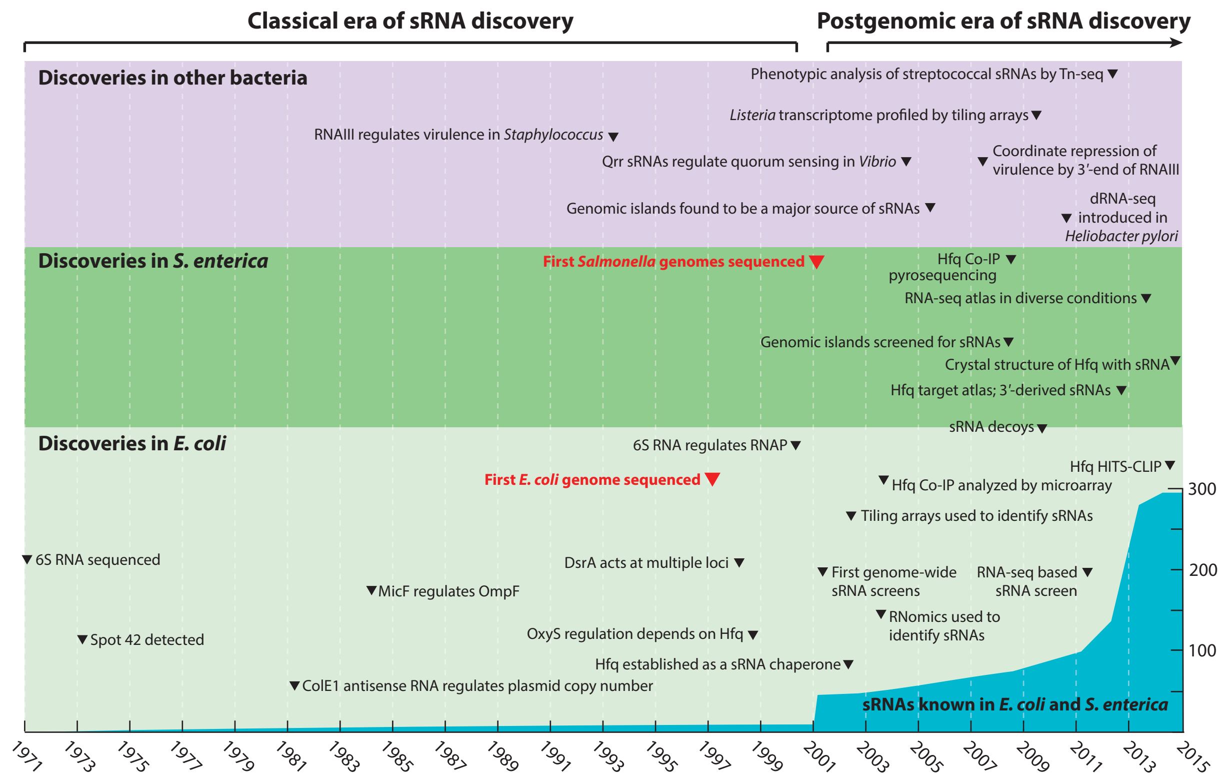

Figure 2

Timeline of sRNA discovery in *Escherichia coli* and *Salmonella enterica*: Timeline of landmark studies and technological innovations in bacterial sRNA biology. The y axis shows the approximate accumulation of sRNA known to be expressed by either *E. coli* or *S. enterica* over time. Studies referenced: 2, 10, 12, 18, 24, 39, 51, 87, 93, 94, 101, 108, 119, 121, 129, 133, 134, 141, 148, 151, 155, 160, 162, 169, 174, 181, 183, 184, 186, 194, 200, 201, 205, 206. Abbreviations: Co-IP, co-immunoprecipitation; HITS-CLIP, high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation.

both screening of targets and validating individual interactions (170). Early screening approaches used transcriptomic or proteomic measurements of strains either overexpressing or deleted for particular sRNAs to identify putative targets. Notwithstanding its success, this approach suffers from the inability to discriminate between direct sRNA targets and downstream effects of, in particular, overexpression. In addition, sustained overexpression can, in some cases, be toxic (1, 168) or cause pleiotropic effects by titrating Hfq (79, 122, 138).

To address these problems, transcriptomic analysis of pulse-expressed sRNAs was introduced (103, 137, 180) and has become the standard. In this approach, an sRNA is briefly expressed from an inducible promoter, followed by transcriptomic analysis to identify rapid changes presumably due to direct sRNA:mRNA interactions and subsequent transcript degradation or stabilization. Regardless of the method used to identify putative targets, individual interactions need to be validated. Verification is typically done through monitoring the effects of compensatory mutations in the sRNA and its target sequence, using a GFP (green fluorescent protein) reporter–based two plasmid system (31, 189) or other reporters, such as the *lacZ* gene (14, 100).

Such approaches are, and will remain, an important cornerstone of sRNA biology. However, even in model strains, these techniques are laborious. If we are to completely characterize the

functional sRNA complement of, say, *E. coli* K-12 or *Salmonella* Typhimurium SL1344, new approaches are needed to prioritize candidates and point out promising directions for research. Additionally, it is increasingly clear that model systems can take us only so far. For example, there is increasing evidence for a diversity of mechanisms that differ between Gram-negative and Gram-positive bacteria, and preliminary work suggests substantial variation in the sRNA complement between species and genera (96). Although the proportion of this variation due to the turnover of non- or weakly functional transcripts is unclear, a number of species-specific sRNAs are clearly functional (134, 148, 186).

Beyond the close relatives of model organisms, the rise of previously obscure bacteria as major opportunistic and nosocomial pathogens, illustrated powerfully by methicillin-resistant *Staphylococcus aureus* (45), epidemic *Clostridium difficile* (71), and recently carbapenem-resistant *Enterobacteriaceae* (179), as well as the adaptation of strains of known pathogens such as *S.* Typhimurium to cause severe disease in at-risk populations (47, 132), provides a clear impetus for the development of technologies that allow us to rapidly characterize sRNAs in nonmodel organisms. The apparent flexibility and rapid evolvability of sRNA-based regulation may enable these molecules to play a general role in the development of these alternative lifestyles. Additionally, disease outcomes are increasingly recognized to be the result of complex interactions between pathogen, host, and host microbiota (150, 190), and recent work has shown some commensals incidentally exacerbate infection (35, 49) or even act as opportunistic cobelligerents (149). Characterizing the diverse organisms of the microbiota and their noncoding genomes will be a major challenge; however, new technologies are emerging that can facilitate this process. In this review, we discuss recent developments in high-throughput sequencing–based approaches with relevance to sRNA biology. Rather than provide a comprehensive list, we use studies selected primarily from *Salmonella* and *E. coli* to illustrate how such generic approaches may accelerate the discovery and functional analysis of sRNAs in any bacterium of interest.

TECHNOLOGIES FOR sRNA CHARACTERIZATION

Until recently, obtaining a global picture of the transcriptional state of a cell has been difficult. Sanger sequencing of cDNA libraries could provide a view of transcripts present in the cell, but the low throughput of the method made genome-wide analyses extremely difficult. Tiling microarrays made quantitative global analyses possible but need to be specifically designed for an organism of interest, have limited resolution owing to a reliance on oligonucleotide hybridization, and are vulnerable to various technical artifacts such as probe mishybridization and saturation. The availability of high-throughput DNA sequencing has driven a revolution in the generation of biological data, allowing the simultaneous interrogation of hundreds to millions of discrete loci through massively parallel DNA analysis. High-throughput technology also enables the related RNA-seq technique wherein RNA is first reverse-transcribed into cDNA and then sequenced. This procedure can be done in ways that preserve information on strandedness and reveal precise boundaries of transcripts. Numerous specialized techniques have been developed based on high-throughput sequencing. Each method has a range of applications (**Figure 3**), spanning a continuum from discovery of sRNAs and annotation of the transcriptome to methods that provide direct insight into sRNA function, allowing us to shape hypotheses that can subsequently be validated through traditional methods.

In the following sections, we address each of these techniques in turn, highlighting their potential applications to sRNA biology. It is important to note that the large amounts of data generated by these methods raises special concerns for both experimental design and analysis compared to traditional molecular techniques; we briefly discuss these issues in two sidebars (see Experimental

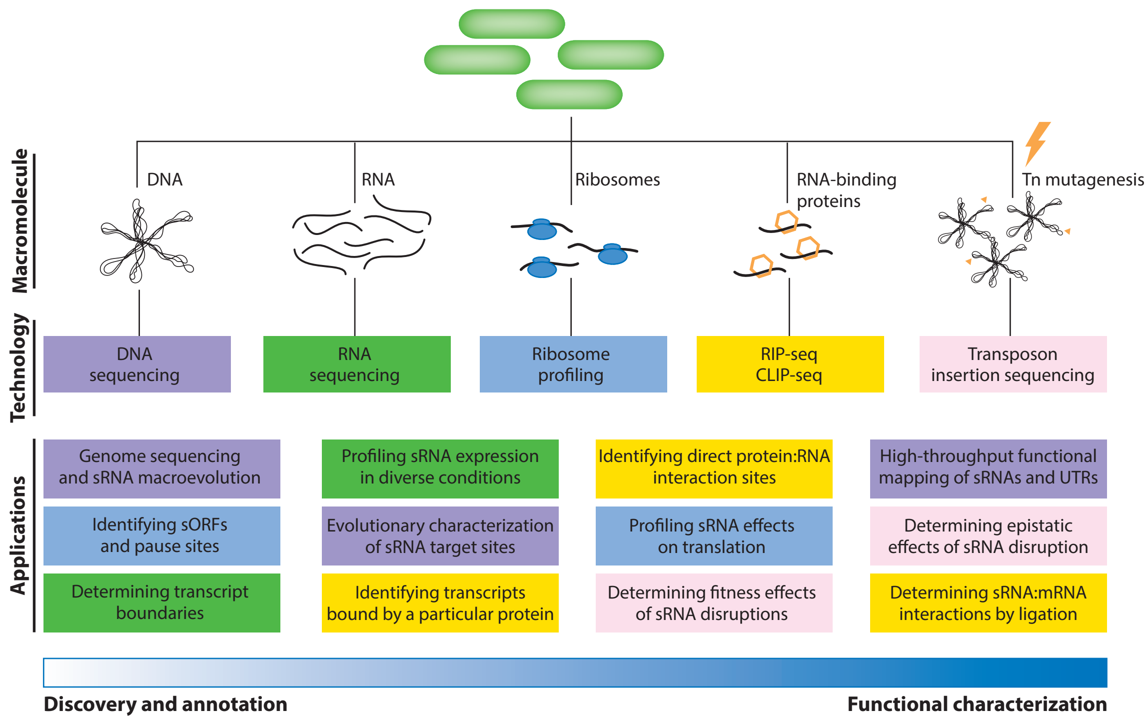

Figure 3

Overview of high-throughput methods for identification and characterization of bacterial sRNAs. Five major technologies are covered in this review. Each is defined primarily by the macromolecule it makes accessible to study: DNA, RNA, ribosomes, RNA-binding proteins, and transposon insertions. The applications for these technologies span a range from discovery of novel transcripts to contribution of direct information about sRNA function. Abbreviations: CLIP, cross-linking and immunoprecipitation; RIP, RNA immunoprecipitation; sORF, short open reading frame; UTRs, untranslated regions.

Concerns for High-Throughput Methods and Analytical Concerns for High-Throughput Methods).

DNA SEQUENCING: sRNA EVOLUTION AND FINE-SCALE FUNCTIONAL MAPPING

Recent studies have revealed substantial variation in sRNAs between related bacteria (96, 143, 182), although the processes governing this variation remain unclear. A number of mechanisms may contribute to this diversity, including horizontal transfer, capture of transcripts spontaneously generated through recombination and mutation, and derivation from coding transcripts (59). Target sites in mRNAs also appear to vary extensively (143, 158), and these alterations may additionally imply variation in sRNA function over evolutionary time. To track the evolution of sRNAs, high-quality collections of genomes from closely related strains are required to determine the precise order of events leading to the birth of novel transcripts and to changes in their regulatory activities.

Studies of genome dynamics in related bacteria are becoming commonplace in epidemiological and evolutionary studies (33, 106). However, such studies generally take a mapping approach (19), in which sequencing reads from a collection of related strains are mapped to a single finished reference genome or pseudomolecule built from assembled fragments, owing to the difficulties in

EXPERIMENTAL CONCERNS FOR HIGH-THROUGHPUT METHODS

Although high-throughput methods have transformed biological research, their appropriate application to different scientific problems is often significantly more complicated than one would expect from reading technical descriptions. Such methods can rapidly produce large amounts of data; however, these data are only as reliable as the experimental procedures used to produce them. Understanding each step of the experimental process and evaluating potential pitfalls when possible are essential to the production of high-quality data. A critical review of CLIP-seq methods (159) provides an excellent example of the analytical approach that should be undertaken in the application of any high-throughput method, documenting known and suspected biases in each step of the protocol, as well as the likely consequences for the interpretation of results. Additionally, as in any experiment, both adequate replication and good controls are necessary. An illustrative example of the importance of well-planned control experiments is given by a recent PAR-CLIP study that used a pull-down of cross-linked FLAG-tagged GFP to demonstrate the prevalence of background cross-linking events (52). By quantifying this background and computationally subtracting it from experiments with genuine RNA-binding proteins, the authors were able to substantially increase the specificity of the assay, providing superior data for downstream applications, such as motif analysis.

assembling even relatively compact bacterial genomes from the short reads delivered by current sequencing platforms (164). Although this approach is often adequate to detect high-confidence single nucleotide polymorphisms for phylogenetic applications or the variable presence of large genomic features, it has difficulties in providing the fine-grain information and structural variation required for detailed functional studies of sRNAs in intergenic regions. Emerging long-read single-molecule platforms, such as PacBio SMRT sequencing (98, 152) or the Oxford Nanopore minION (5, 153), are beginning to provide a solution by simplifying the assembly problem (87).

A recent study in *Helicobacter pylori* provides an excellent example of how the availability of closely related genome sequences can provide insight into the evolution of sRNA regulation (145). The authors characterized the sRNA RepG, a highly abundant sRNA discovered during global transcriptional mapping of *H. pylori* (169). They found that in *H. pylori* strain 26695, RepG

ANALYTICAL CONCERNS FOR HIGH-THROUGHPUT METHODS

High-throughput methods raise special concerns regarding analysis. Measuring large numbers of loci simultaneously greatly increases the chances of spurious results. The analysis of such data must be done carefully and can often be as time consuming as the development of the assay itself. Issues such as data normalization, distributional assumptions made by common statistical tests, and appropriate corrections for multiple testing must be considered in light of the specific methods used. For commonly applied methods, such as comparing conventional RNA-seq data sets, standard analytical tools are available and can be applied with a minimal computational background (92, 99, 107). However, quality control and assessment of reproducibility are essential. For many emerging technologies, the best analysis techniques remain open to debate. Although it may be impractical for every biologist to additionally become a programmer and a statistician, cultivating an appreciation for the subtleties of large data sets, experimental design, and statistics is an important part of interpreting results. Major journals have begun running series of articles aimed at introducing experimentalists to statistical analysis (126) and programming tools (125). It is important that bioinformaticians also familiarize themselves with experimental protocols and have input into the design of experiments so that potential sources of bias can be effectively accounted for in analysis.

regulates the *tlpB* mRNA encoding a chemotaxis receptor: a C/U-rich sequence contained in a hairpin loop of RepG base pairs with a G-repeat located in the 5′ UTR (untranslated region) of *tlpB* (145). As homopolymeric stretches are known to be highly variable in related bacteria, even responsible for phase variation through polymerase slippage during replication (124), the authors also examined the impact of G-repeat length on regulation. Optimal regulation occurred in a narrow range of G-repeat lengths, with shorter or longer repeats abrogating regulation by RepG. The evolutionary relevance of the relationship between repeat length and gene regulation was established by examining the *tlpB* locus in 32 *Helicobacter* strains, demonstrating wide variation in G-repeat length in naturally occurring strains (145).

Although the above study focused on only a single target region, it illustrates how phylogenetically dense genome sequencing can shed light on the functional evolution of sRNAs. As more genomes of large strain collections are sequenced, such evolutionary analysis of target regions may become a routine part of sRNA characterization. A version of this approach has been taken by the CopraRNA target prediction method, which uses a measure of the evolutionary conservation of mRNA target sites to improve predictions (203).

High-throughput DNA sequencing can also be used to directly derive functional information about sRNA:target interactions through monitoring saturated libraries of point mutants (77, 146). Traditionally, validation of target sites has been performed by monitoring reporter activity following disruptive mutations at predicted intermolecular base pairs, which should abolish regulation, and compensatory mutations, which should restore near wild-type regulation (170). Although two-plasmid systems with, for example, a fluorescent GFP reporter (31, 189) streamlined these experiments, they remain labor intensive and often must be repeated numerous times, as computational RNA:RNA hybridization tools only roughly estimate the contribution of individual nucleotides to an interaction. However, a high-throughput adaptation of this approach now allows every nucleotide of a regulatory sequence of interest to be assayed simultaneously in a single experiment using a library of mutagenized target variants (77). The effects of sequence variants on protein production can then be monitored through specific sequencing of the mutagenized regions in fluorescence-activated cell sorting (FACS)-derived fractions of bacteria containing a two-plasmid system, similar to that used in traditional interaction validation experiments. This has been applied to the 109-base 5′ UTR of the *csgD* gene in order to map the regulatory region for the OmrA sRNA (77). By sorting these variants into high, wild-type, and low GFP fractions before sequencing, the authors identified mutations in the previously known OmrA-binding site as having strong effects on GFP expression in the presence of OmrA. Interestingly, mutations in the structurally accessible region of the target sequence were clearly enriched in the high-GFP fraction compared with those in the base-paired region, adding additional experimental evidence to the claim that accessibility is a major determinant of sRNA binding (9, 142, 158).

The reciprocal approach, i.e., FACS of mutagenized sRNAs and wild-type target-GFP fusions, has also been taken (146) (**Figure 4*a***). Two Hfq-dependent sRNAs, RyhB and DsrA, were mutagenized and expressed in the presence of two known target UTRs each. This report contains a number of interesting technical innovations. For instance, to counteract the inherent noise in FACS sorting, bacteria are sorted into six fractions, and then the resulting distribution of read counts across these fractions is used to estimate the GFP level of a variant. These measurements provide additional experimental evidence for two commonly held beliefs about sRNA-based regulation: that the strength of target regulation can be tuned through mutations that affect the thermodynamics of sRNA:target hybridization, and that sRNA structural stability contributes to regulation efficiency, possibly through promotion of transcript abundance. An analysis of genetic interactions in double mutants present in the library revealed expected patterns of covariation in stem structures. However, this analysis also identified hub mutations, which appeared to interact

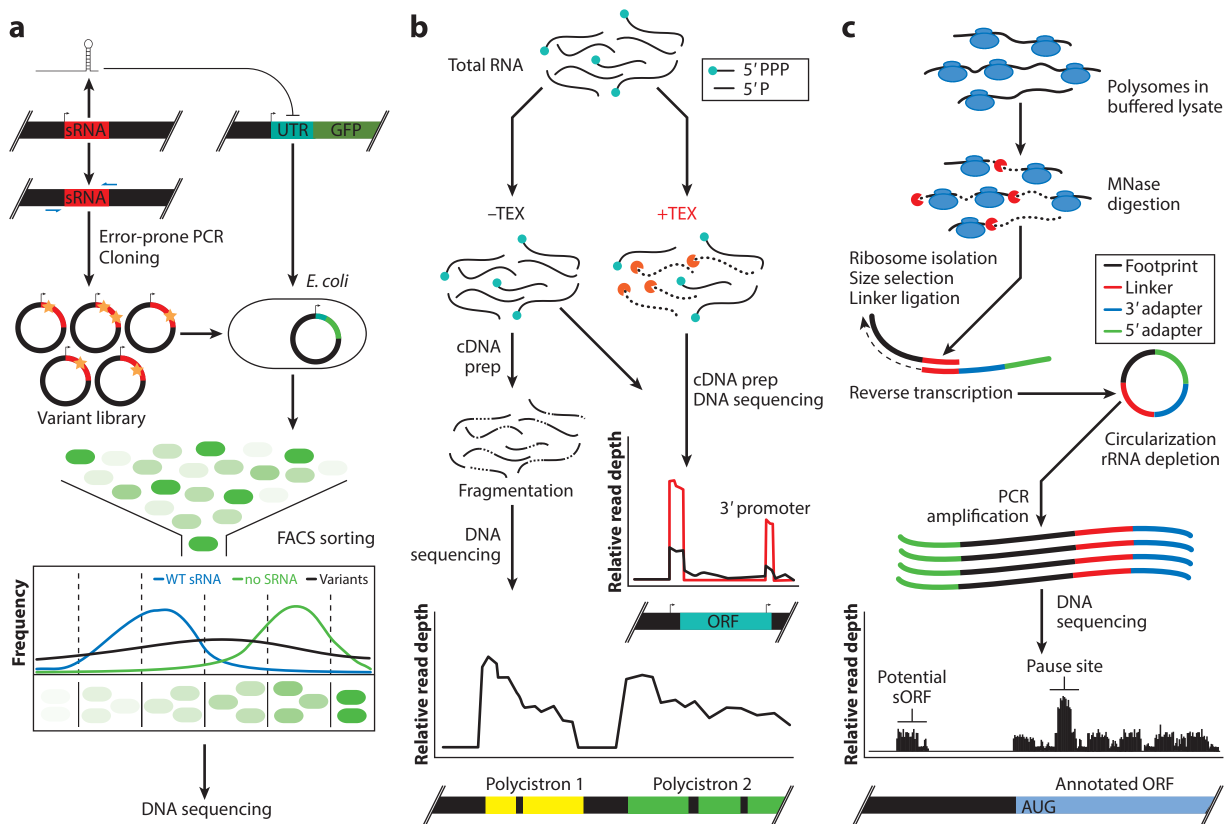

Figure 4

High-throughput point mutagenesis, RNA-seq, and ribosome profiling. (*a*) Overview of the method for monitoring sRNA variant libraries described in Reference 146. A large number of sequence variants are generated through error-prone polymerase chain reaction (PCR) of an sRNA of interest. These variants are then cloned to generate a plasmid expression library. This library is transferred by electroporation into *Escherichia coli* carrying a second plasmid containing the untranslated region of a known target gene fused to a fluorescent reporter. The resulting population of bacteria is sorted using FACS (fluorescence-activated cell sorting), and DNA sequencing of the sorted fractions is used to determine variant effects on protein expression. Adapted from Reference 146. (*b*) Conventional and differential RNA-seq. Total RNA is isolated from a population of bacteria in a condition of interest. dRNA-seq (171) (*right path*) proceeds by dividing the RNA into two fractions. One fraction is treated with TEX (5′ P-dependent terminator exonuclease), which specifically degrades processed transcripts lacking a 5′PPP (5′ triphosphate), prior to conversion to cDNA and sequencing, while the other fraction is not treated. By examining the differences in read enrichment between these two libraries, both native and processed 5′ transcript boundaries can be determined. Conventional RNA-seq (*left path*) proceeds by the direct conversion of total RNA to cDNA and fragmentation before sequencing. This technique provides information on relative transcript abundance and can be used to help define polycistronic structures. (*c*) Ribosome profiling, as described in Reference 131. Cell lysates are subjected to micrococcal nuclease (MNase) digestion, leaving discrete ∼30-base footprints physically enclosed by the ribosome, before intact ribosome-footprint complexes are recovered by sedimentation. A specialized circularization process is used to minimize biases in library preparation from short RNA fragments (81), and rRNA-derived sequences are depleted before library amplification and sequencing. Mapping of the resulting reads provides information on ribosome occupancy and can be used to discover unannotated open reading frames and pause sites, as well as the relative rates of translation. Abbreviations: GFP, green fluorescent protein; PCR, polymerase chain reaction; UTR, untranslated region; WT, wild-type.

with many other positions within the sRNA, including within the seed region. Of particular interest was the U55A mutation in RyhB, which appears to relax the secondary structure, allowing for continued regulation of target sequences even in the presence of mutations to the seed sequence that would otherwise have abrogated sRNA function (146). These observations suggest a complex relationship between sRNA secondary structure, target specificity, and regulation efficiency, and provide a starting point for the interpretation and functional analysis of naturally occurring sRNA variants.

RNA SEQUENCING: AIDING sRNA DISCOVERY AND FUNCTIONAL ANALYSIS

sRNAs operate through direct base-pairing interactions with mRNA or other transcripts, often changing the expression levels of these cellular targets. Therefore, global methods to determine potential expression level changes of all the individual transcripts of a cell are foundational to understanding how an sRNA might act. RNA-seq is rapidly replacing microarray-based gene expression analyses by obviating the need for custom microarray design and by providing increased sensitivity and dynamic range, while reducing technical artifacts such as nonspecific probe hybridization (34). In contrast to microarrays, RNA-seq allows the high-resolution assay of transcriptional changes independently of annotation. This strategy seems particularly important given recent reports of sRNA interactions with other noncoding transcripts (90, 117, 186). The high sensitivity of RNA-seq can also enable previously difficult or impossible applications, such as simultaneous transcriptomic analysis of intracellular pathogens and their hosts, as proposed in Reference 202, or single-cell transcriptomics (163). Here, we highlight one transformative potential of RNA-seq: its ability to map the structural complexities of the bacterial transcriptome, a prerequisite for a complete determination of sRNA interactions.

Assuming that most sRNAs act by base-pairing mechanisms and can operate at 5′ UTRs and also within coding sequences, 3′ UTRs, and the intergenic spaces between polycistronic genes, high-resolution transcriptome structures are key to gaining a full understanding of sRNA function. However, even the simplest bacteria appear to possess a remarkably complicated transcript complement (63, 70), which varies extensively between related strains and species (40, 83, 96, 197). Bacterial genes are frequently cotranscribed in polycistronic structures, and alternative, cryptic, and antisense promoters now appear common. Indeed, several cases are already known in which an sRNA acts to differentially regulate polycistronic genes, either through inducing degradation of a portion of the polycistron message (38), stimulation of premature transcription termination (15), or selective stabilization of a particular isoform (53, 139). A clear example is given by the SgrS:*yigL* interaction in *Salmonella* that regulates glucose homeostasis, in which SgrS acts on the 5′ end of the monocistronic form of the *yigL* mRNA generated by an RNase E cleavage within an upstream protein-coding sequence (139). It is unlikely that this interaction could have been understood in the absence of a transcriptional atlas of *Salmonella* (88). Producing such an atlas requires reliable predictions of transcript boundaries, which can be provided through specialized RNA-seq protocols.

A number of techniques for global mapping of the 5′ ends of primary transcripts have been developed (29, 112, 169, 173, 204). One approach in particular, called differential RNA-seq (dRNA-seq) (169) has provided insight into the transcriptome structures of a wide range of bacteria (reviewed in 171) (**Figure 4*b***). This method relies on the fact that primary bacterial transcripts contain a 5′ triphosphate (5′PPP), whereas the 5′PPP is removed from processed transcripts. These two classes of transcripts can be discriminated by 5′P-dependent terminator exonuclease (TEX), which selectively degrades monophosphorylated processed transcripts. In addition to mapping

5′ ends of mRNAs and intergenic transcripts, this technique can particularly aid in the discovery of an emerging class of sRNAs derived from the 3′ regions of mRNAs, which can be generated either by independent promoters or through cleavage of the primary coding transcript (24, 118). Such RNAs are difficult to identify with conventional transcriptomics.

A trailblazing application of dRNA-seq has been its use in constructing an infection-relevant atlas of the *S.* Typhimurium transcriptome (87). The authors first performed conventional RNA-seq under 22 conditions relevant to the study of *S.* Typhimurium's pathogenic lifestyle. These conditions included a variety of growth phases in standard laboratory conditions: growth in in vitro conditions that replicate aspects of environments that *S.* Typhimurium may encounter in vivo, such as anaerobic conditions relevant to portions of the intestinal phase of infection; growth in low phosphate, low magnesium minimal media that replicates aspects of the intracellular environment; and, finally, exposure to a variety of shocks that *Salmonella* must survive to establish a successful infection, including iron limitation, varying temperatures, acidic conditions, and bile. Next, dRNA-seq was applied to four of the in vitro conditions and a pool composed of RNA from all 22 conditions. By comparing the results obtained from TSSs (transcription start sites) predicted by dRNA-seq of the pooled sample to those from individual conditions, the authors showed that dRNA-seq of pooled RNA can identify 96% of TSSs predicted by dRNA-seq of individual conditions, and the pooled dRNA-seq data successfully predicted TSSs for genes that were only expressed in 1 of the 22 assayed conditions by conventional RNA-seq. However, because the potential biases of pooling dRNA-seq libraries are unclear, multiplexing of samples provides a cleaner alternative to pooling (87). Nevertheless, this study should inspire the construction of similar atlases in a diverse range of microbial species and growth conditions. Such atlases provide valuable information about the conditions under which particular sRNAs and their targets are expressed and, hence, in which their activities should be investigated.

Until recently, less attention has been given to the determination of the 3′ ends of transcripts and their roles in regulation in bacteria. Although the sequencing of fragmented RNA provides an estimate of 3′ end positions (32), the signals resulting from this technology are often unclear, perhaps owing to stochastically driven sloppy transcriptional termination in some bacteria (3). Accurate methods for prediction of Rho-independent intrinsic terminators are available (55) and can provide some insight into transcript boundaries. However, read-through at intrinsic transcriptional terminators is frequent (22, 28, 178), may play a functional role in affecting the stoichiometry of operon components, and can even be differentially regulated in response to stimuli (4). These effects can be estimated using RNA-seq (30). Additionally, Rho-dependent termination varies across the bacterial phylogeny and these termination sites can be difficult to predict computationally (147). High-throughput sequencing-based 3′ RACE (rapid amplification of cDNA ends) protocols have been applied to individual genes of interest (155) and may provide a solution to defining 3′ transcript ends, although the lack of a clear signal, similar to the 5′PPP used in dRNA-seq protocols for mapping 5′ ends, makes discriminating native ends from the products of processing and degradation challenging.

Once transcript end points have been determined, the major problem remains in how to use them to determine polycistron structure and the relative usage of alternative transcripts. Recent comparative studies of TSSs in related bacterial species suggest that the transcript variants produced by a locus can differ significantly, even on short timescales (40, 83). Although computational methods exist for predicting operons and show reasonable accuracy using the incomplete experimentally confirmed data sets available (17), most of these algorithms do not take transcriptomic data into consideration and therefore are silent on the lengths of 5′ and 3′ UTRs, relative transcript usage, and the existence of suboperons and alternative transcripts. Recent methods have been proposed for using information from fragmented RNA-seq data to serve as a link between

adjacent genes in the prediction of polycistronic transcripts (102, 109). These methods show good performance, although again, they do not address the existence of alternative transcripts. Ultimately, the problem of alternative transcript prediction and usage quantification is likely to be difficult because of the nonuniformity commonly observed in short-read RNA-seq data, as well as underlying stochasticity in transcription and processing. However, the challenge of analyzing alternative transcripts bears some similarity to the more well-studied problem of alternative isoform utilization in eukaryotes (66, 82, 154, 185, 199), and it is possible that techniques developed to address the latter problem may be adaptable to bacteria. The ability to monitor differential isoform usage provides a layer of information previously unavailable in microarray or conventional RNA-seq experiments, and should help to identify additional sRNAs that operate through the differential regulation of cistrons within transcriptional units (34, 13, 47, 127).

RIBOSOME PROFILING: TRANSLATIONAL DYNAMICS

Although many sRNAs act, at least in part, by changing transcript levels, the effect of sRNA regulation on phenotype ultimately depends on changes in protein levels which, however, can also occur by a primary translational regulation without a strong associated change in the target transcript level. Thus, a purely transcriptomic approach to characterizing sRNA-based regulation may miss important interactions. Directly monitoring sRNA-induced changes in protein abundance through proteomics has been successful (41, 68, 156, 168, 188), although such methods are labor intensive and difficult to apply in unbiased, genome-wide fashions. By contrast, the RNA-seq-based ribosome profiling technique captures translational dynamics, in theory, of all cellular proteins (80), and thus provides a bridging technology between traditional RNA-seq experiments and the direct measurement of protein levels through, for instance, mass spectrometry.

Like many recent high-throughput techniques, ribosome profiling adapts an old technology, ribosome footprinting, to take advantage of the massively parallel measurements afforded by modern short-read sequencing (**Figure 4*c***) (80, 81). Cycloheximide treatment [in eukaryotes (81)], chloramphenicol treatment [in bacteria (131)] or rapid filtration (95, 131) and flash freezing are used to halt translational elongation. Polysomes are treated with RNase, leaving short ~30-base footprints physically enclosed by the translating ribosome. These RNA fragments are subsequently used as templates for the production of cDNA, which can be analyzed through deep sequencing as in conventional RNA-seq. The relative abundance of footprints covering a particular region is indicative of ribosome occupancy at that region, providing a snapshot of ongoing protein synthesis in the cell.

Ribosome profiling has recently been applied to bacterial sRNAs (65). In characterizing an sRNA derived from the 3′ UTR of the *cutC* gene in *E. coli*, dubbed MicL, Guo and colleagues found through a pulse induction approach coupled with RNA-seq that this sRNA regulates the outer membrane lipoprotein Lpp. Although this was the sole target identified by transcriptomics, additional targets regulated solely at the translational level could not be ruled out. By applying ribosome profiling to similarly treated samples, the authors demonstrated that the *lpp* mRNA is the primary, if not the only, target of MicL. Finally, the authors showed that the copper sensitivity phenotype previously attributed to disruption of CutC is entirely due to MicL and its regulation of Lpp expression, illustrating the importance of understanding transcriptional complexity for the interpretation of phenotype. Although in this proof of principle case, ribosome profiling primarily served to reinforce results derived from traditional transcriptomics, clear cases are known in which translational regulation is not accompanied by transcript degradation and therefore would not be detected by transcriptomic approaches such as RNA-seq (76, 78). The application of ribosomal profiling to additional sRNAs is likely to reveal novel targets and regulatory mechanisms, as well

as provide insights into the dynamics and relative importance of transcriptional degradation and translation repression in sRNA-based regulation.

INVESTIGATING PROTEIN PARTNERS: RIP-SEQ AND CLIP-SEQ

As discussed, many bacterial sRNAs are at least partially dependent on protein-binding partners for their functions, Hfq being a prominent example, with much of the enterobacterial sRNA complement requiring it for function (25, 37, 195). Immunoprecipitation of an RNA-binding protein (RBP) of interest followed by isolation and analysis of bound RNA [RIP (RNA immunoprecipitation)] allows rapid characterization of its RNA regulon (**Figure 5*a***). The earliest applications of this technique to sRNAs in bacteria used microarrays carrying a range of genic and intergenic probes to explore the Hfq-binding landscape of *E. coli* and identified a number of novel Hfq-binding sRNAs, as well as provided evidence for widespread binding of mRNA transcripts and a role for Hfq in tRNA processing (206). By adapting this technique to utilize high-throughput pyrosequencing in place of microarrays, Sittka and colleagues probed the *Salmonella* genome for Hfq-binding transcripts, more than doubling the number of known Hfq-binding sRNAs and providing evidence for a major role for Hfq in regulating horizontally transferred sequences (174). Subsequently, Illumina sequencing enabled comprehensive profiling of the Hfq interactome throughout growth to identify as many as 25% of *S.* Typhimurium's mRNAs as members of the Hfq regulon and revealed large changes in the sRNA complement bound by Hfq during transitions in growth phase (24). Such a RIP-seq-based atlas of Hfq-bound transcripts, containing both mRNAs and sRNAs, is an excellent starting point for sRNA research in an organism of interest (24) and may pinpoint functional RNA species among seeming degradation intermediates. For example, a recent study has identified that the SroC RNA, an apparent decay product of a polycistronic mRNA encoding an ABC transporter that was strongly bound by Hfq (24), acts as an antagonist of the sRNA GcvB (117).

Although RIP-seq has proven its utility in these and other studies, it suffers from technical limitations. First, it is limited to the detection of stable protein:RNA interactions that can survive the immunoprecipitation procedure, although there is also a potential to detect nonspecific interactions because of the gentle handling required and for the formation of spurious associations following cell lysis (116). Additionally, as complete or only partially fragmented or degraded transcripts are pulled down by this method, the resolution of RIP-seq is poor, and it is unclear whether it can provide reliable information on the locations of RBP binding sites within a transcript. These issues make certain downstream analyses, such as the identification of RBP binding motifs, difficult.

Modifications of RIP-seq protocols incorporating direct in vivo UV cross-linking between RBPs and their RNA targets (**Figure 5*a***), commonly referred to as HITS-CLIP (high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation) or CLIP-seq, have been driving a revolution in the study of eukaryotic RBPs (36, 84, 110, 115, 157). The physical cross-linking of binding partners allows the use of more stringent washes, eliminating nonspecific interactions. Additionally, RNase treatment can be used to increase the resolution of the assay compared to RIP-seq. Refinements of this protocol, known as PAR-CLIP and iCLIP, introduce additional modification that can further improve the ability to detect direct interaction sites at nucleotide resolution (67, 85). It should be noted that not all RBP:RNA interactions are UV cross-linked with the same efficiency, and some true interactions, particularly those not involving direct interactions with a nucleobase, will certainly be missed by this method, although the factors influencing these biases are presently unclear (110).

HITS-CLIP has recently been applied in a bacterial system for the first time, to study Hfq binding in a strain of enterohemorrhagic *E. coli* (EHEC) (186). In addition to discovering numerous

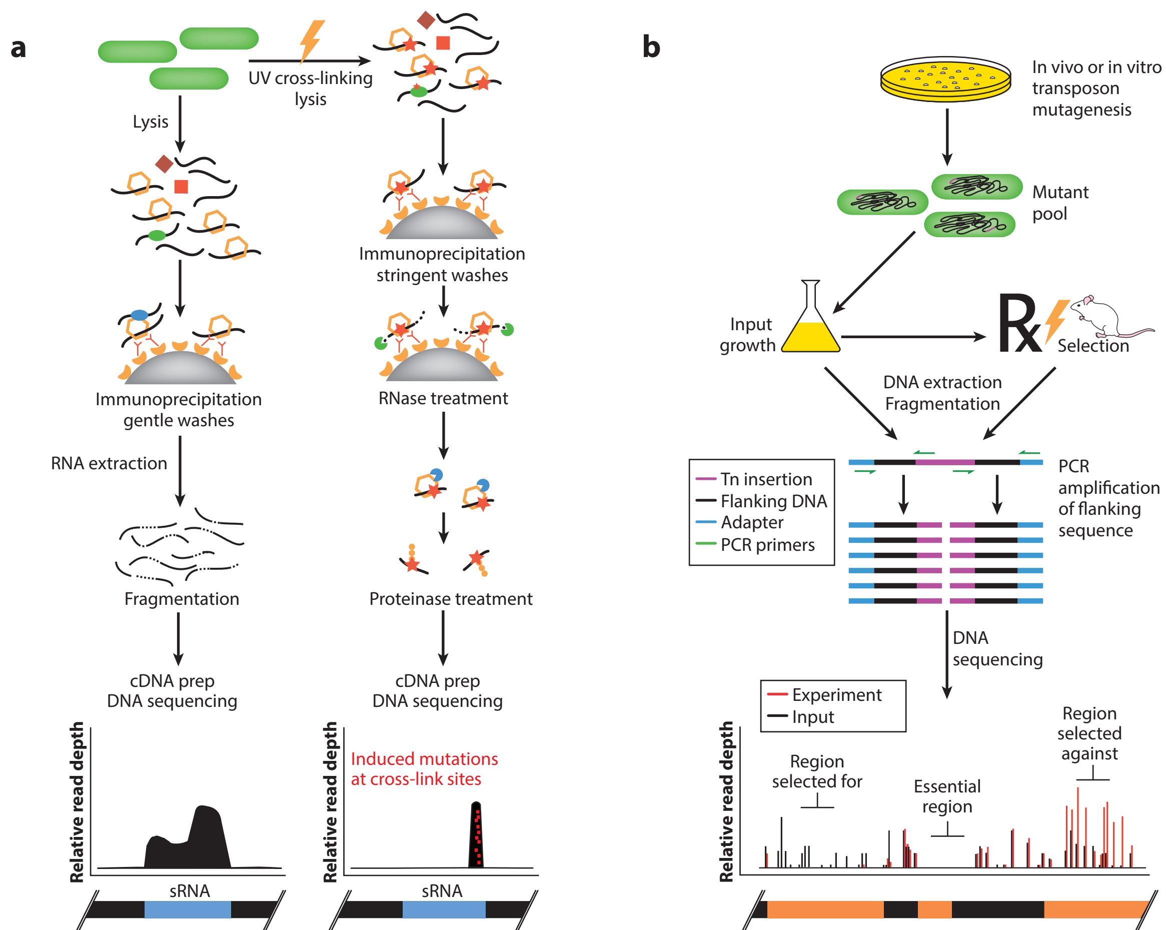

Figure 5

RIP-seq, CLIP-seq, and transposon insertion sequencing. (*a*) Comparison of RIP-seq and CLIP-seq protocols. In RIP-seq (*left path*), an RNA-binding protein (RBP) of interest is immunoprecipitated on streptavidin beads following cell lysis. As the RNA is captured through relatively weak RNA:protein interactions, care must be taken in washing the sample so as not to disrupt these interactions. However, this gentle handling can lead to the additional capture of RNAs interacting indirectly with the RBP, either through RNA:RNA interactions or protein:protein interactions. In CLIP (*right path*), cells are subjected to UV cross-linking prior to lysis, introducing covalent cross-links between proteins and proximal nucleobases. This strategy allows more stringent washing and RNase trimming of the bound RNAs, increasing the specificity of the assay and the resolution of recovered binding sites. Proteinase K digestion is then used to digest the RBP, leaving a short polypeptide at the cross-link site, which may induce polymorphisms during reverse transcription. (*b*) Generic overview of transposon insertion sequencing protocols. A large collection of transposon mutants is generated, generally through the use of a Tn*5* or *mariner*-derived transposon, and pooled. This mutant pool is then briefly expanded in vitro to attain sufficient bacterial numbers for the assay of interest. The resulting population is subjected to some form of selection, for example growth in an infection model, exposure to a drug, or survival during stress conditions. DNA is extracted from both the input pool and the bacteria surviving the assay, and transposon-flanking sequences are PCR (polymerase chain reaction) amplified using transposon-specific primers. The relative contribution of a region of interest to fitness can be estimated from the ratio of sequencing reads recovered from the input and the experimental conditions. Adapted from Reference 7.

island-encoded Hfq-binding sRNAs, the high resolution of the HITS-CLIP assay allowed the authors to study characteristics of the Hfq-binding regions of RNAs. Among other findings, this study confirmed the previous observation that sRNA target sites on mRNAs are often in close proximity to Hfq-binding sites (9), suggesting that the availability of high-resolution Hfq:mRNA binding site data may be one avenue by which to improve computational predictions of sRNA targets. Finally, in depth analysis of the novel sRNAs discovered through this method identified two so-called anti-sRNAs encoded within prophage elements (186).

UV irradiation has been the most popular method for performing CLIP because of its high specificity for direct RNA:protein interactions; however, many alternative cross-linking treatments exist with varying specificities and efficiencies, and can be applied in CLIP protocols to provide additional types of information. Examples include formaldehyde, which cross-links both RNA:protein and protein:protein interactions, identifying both direct and indirect interactions, and 4′-aminomethyltrioxalen, which induces cross-links between proximal uridine bases, providing evidence for direct RNA:RNA interactions (44). Careful application of suites of cross-linking treatments with varying properties to samples of interest, coupled with differential analyses, can allow the high-throughput mapping of layers of RNA:RNA, RNA:protein, and indirect protein:protein-mediated interactions. These approaches have been applied in eukaryotic systems for both the detailed study of the interactions of individual noncoding RNAs (ncRNAs) (44), and the genome-wide mapping of RNA:protein interaction sites (57, 172). UV CLIP can also be modified to provide direct evidence for ncRNA:target interactions, through the introduction of a ligation step during RNA isolation in CLASH (cross-linking, ligation, and sequencing of hybrids)-like protocols (89). This method generates chimeric ncRNA:target reads and has been successfully applied to the identification of noncanonical targets for both snoRNAs (small nucleolar RNAs) (89) and microRNAs (62, 73) in eukaryotic systems. The adaptation of such a protocol to sRNAs would provide a high-throughput method for directly assaying sRNA:mRNA interactions without relying on secondary readouts (transcript or protein levels), and promises to unveil interactions invisible to current approaches.

TRANSPOSON INSERTION SEQUENCING: FITNESS TO FUNCTION

Although monitoring the transcriptional and translational outputs of the cell is informative about the processes occurring within the cell, taken on their own, they are effectively silent regarding the relative importance of these processes to survival in any particular environment. Transposon mutagenesis allows the direct monitoring of the fitness effects of gene disruptions and has underpinned numerous insights into the biology of important human pathogens, including *S.* Typhimurium, *Mycobacterium tuberculosis*, and *Shigella flexneri* (105). Transposon insertion sequencing (TIS) (recently reviewed in 7, 193; see also 101 for an excellent early application to sRNAs) provides a high-throughput method to quantify the effects of these disruptions (**Figure 5*b***). A number of protocols have been developed to serve this purpose, including TraDIS (transposon-directed insertion site sequencing) (91) and Tn-seq (transposon sequencing) (191), with variations in technical details. These techniques allow monitoring of large single insertion mutant libraries by sequencing. These libraries can be grown in arbitrary conditions, including many animal models of disease, and changes in read frequency after growth serve as a proxy for changes in mutant abundance. As transposition is effectively random, all regions of the genome, including noncoding regions (7), can be queried given sufficient insertion density, although the bottlenecks present under in vivo conditions can limit feasible pool complexity (26, 60, 104).

A recent study with implications for sRNA biology has applied TIS to globally identify *S.* Typhimurium genes involved in infection of four animal models, those of mouse, chicken,

swine, and cattle (26). The murine typhoid model is a widely used proxy for human typhoid fever (187), whereas infection of food animals mimics aspects of human gastrointestinal disease and serves as an important disease reservoir (176). The analysis of a library of 7,702 (9,792 in mouse) *S.* Typhimurium mutants passed through these four animal models recovered known virulence determinants and suggested a variety of determinants of host specificity. In a single set of experiments, this analysis also identified 11 sRNAs as having potential virulence phenotypes, 8 of them in 2 or more independent infection models, more than tripling (72) the number of *Salmonella* sRNAs with potential roles in pathogenicity. However, the ability to survey short genomic regions depends on insertion density; thus, this study (27) was able to assay only 40 of the 140 *Salmonella* sRNAs known at the time, and many of those through only a single insertion in each.

Bacterial sRNA disruptions generally show mild phenotypes, and this may be due to the existence of redundant pathways providing a fail-safe mechanism for stress responses in which sRNAs are frequently involved. An additional feature of post-transcriptional regulation is that it is inherently epistatic (58), i.e., the effect of an sRNA regulator depends on the transcription factors active in the cell under the particular condition assayed and so the output has the potential to vary dramatically. Although large target regulons have been identified for many sRNAs, it is often unclear which targets have important effects on physiology in any particular environment and hence which drive survival under biologically relevant conditions. Understanding these interactions requires tools that can measure the sRNA impact on organismal fitness, a role for which TIS is well suited. One of the first applications of TIS was to measure the epistatic effects of disruption of two discrete loci through the generation and analysis of transposon mutant libraries in defined deletion backgrounds, mapping genetic interaction networks for several transcriptional regulators and ABC transporters (191).

TIS has recently been applied to the bacterial ncRNA tmRNA, involved in freeing stalled ribosomes, in *Caulobacter crescentus*. The technique successfully identified synthetic lethal interactions with a novel protein-based ribosome release factor (46), specifically demonstrating the power of this approach in identifying redundant pathways in stress responses. Applying similar methods in sRNA deletion backgrounds under a variety of physiologically relevant conditions could provide an avenue for moving from simply mapping post-transcriptional regulation toward an understanding of consequences for fitness, as well as for dissecting the effects of the complex network architectures common in many processes regulated by sRNAs (8).

TOWARD AN INTEGRATIVE VIEW OF sRNA FUNCTION

One of the main features of many of the high-throughput, postgenomic methods reviewed here is that they can produce data sets with implications potentially far beyond the initial questions they were meant to address (69). In fact, a number of the studies highlighted here have explicitly set out to generate high-quality reference data sets. Other notable examples of this genus of study include comprehensive multi-omics annotation of the *C. crescentus* genome (167) and phenotype-genotype mapping in a wide variety of in vitro and in vivo conditions using TIS in *Streptococcus pneumoniae* (192). Even in cases in which data are not generated with the purpose of providing a reference, the genome-wide measurements produced can often be reused for unrelated studies. The potential for applications to sRNA biology is shown by studies integrating transcriptomic data from a variety of sources for the study of sRNA evolution (96) and interaction networks (120).

To illustrate the power of integrating postgenomic data sets for the characterization of sRNAs, we have performed a simple reanalysis of five large data sets touching on the role of sRNAs in *Salmonella* virulence (**Figure 6**). As discussed briefly above, *S.* Typhimurium is a major pathogen

causing primarily a self-limiting gastroenteritis in humans, although some strains are capable of causing invasive systemic disease in susceptible, primarily immunocompromised, populations (47). Many animal models exhibit disease symptoms reminiscent of human gastroenteritis; however, susceptible mouse strains develop diseases reminiscent of human typhoid fever, presenting either an acute or a controlled systemic disease, depending on host genotype (187). Importantly, these various models of disease can differ in the virulence factors required to successfully establish an infection (166). If our goal then is to understand the contributions of sRNAs to virulence programs, it is important to know which disease model is most appropriate and at what stage of infection and under which conditions the sRNA may be active. Although our example here is specific to *S.* Typhimurium, these general points could apply to any bacterium in any model system.

A fairly minimal analysis, presented in **Figure 6**, already provides a number of insights into the potential function of sRNAs in infection. Thirty-eight sRNAs have now been predicted by at least one in vivo fitness challenge to give measurable virulence phenotypes, with 21 of these findings supported by at least 2 independent animal models, a nearly eightfold increase from those suspected only a few years ago (72). We can additionally use this table to begin constructing specific, testable hypotheses. For instance, InvR has been characterized as a pathogenicity island–encoded sRNA regulating the translation of the major outer membrane porin OmpD (148). Even though it was shown to be coregulated with the pathogenicity island SPI-1, no virulence phenotype was reported at the time. Although the deletion of this sRNA showed no phenotype in a systemic mouse model of infection (165), TIS assays in three gastrointestinal models of infection indicate an attenuation of virulence. Additionally, although transcriptional profiling in these 14 selected conditions show strong constitutive expression of InvR, it appears to be specifically induced in anaerobic growth and under oxygen shock, in addition to its previously reported accumulation in the early stationary phase, further indicating a role in survival in the gastrointestinal tract. Although these predictions remain to be validated, they provide a view of how the existence and integration of such data sets can open a window into sRNA function and guide directions for molecular studies.

PERSPECTIVE

Bacterial sRNAs are pervasive and appear to be a universal feature of bacterial transcriptomes. Modern transcriptomics techniques have made the discovery of novel transcripts straightforward. However, established methods for functional characterization simply cannot keep pace with the increasing numbers of noncoding transcripts being discovered in the increasingly diverse menagerie of bacteria being studied for scientific, medical, and industrial purposes. In this review, we have highlighted a number of emerging techniques based on high-throughput sequencing that can provide rapid insight into sRNA function. Many sRNAs appear to be modulators of expression rather than serving as simple on-off switches, and they have subtle effects on phenotype. Understanding the implications of these subtle regulatory roles requires using global approaches to explore the effects of this tuning of the transcriptome on both gene expression and fitness across varied conditions and dynamically during transitions. Other technologies will also prove useful in this realm, such as high-throughput phenotyping (128), sequencing-based approaches to RNA secondary structures (42, 144), and direct pull-down and sequencing of sRNA targets (90). The increasing accessibility of high-throughput methods provides us with an opportunity to begin systematically exploring the interactions between sRNAs, the transcriptome, and the environments in which bacteria reside. For example, simultaneous high-throughput sequencing of the transcriptomes of a pathogenic bacterium and its host (202) promises to reveal molecular phenotypes of sRNA deletions that manifest in altered bacterial stress responses or host defense programs without evident macroscopic changes within the time of the assay.

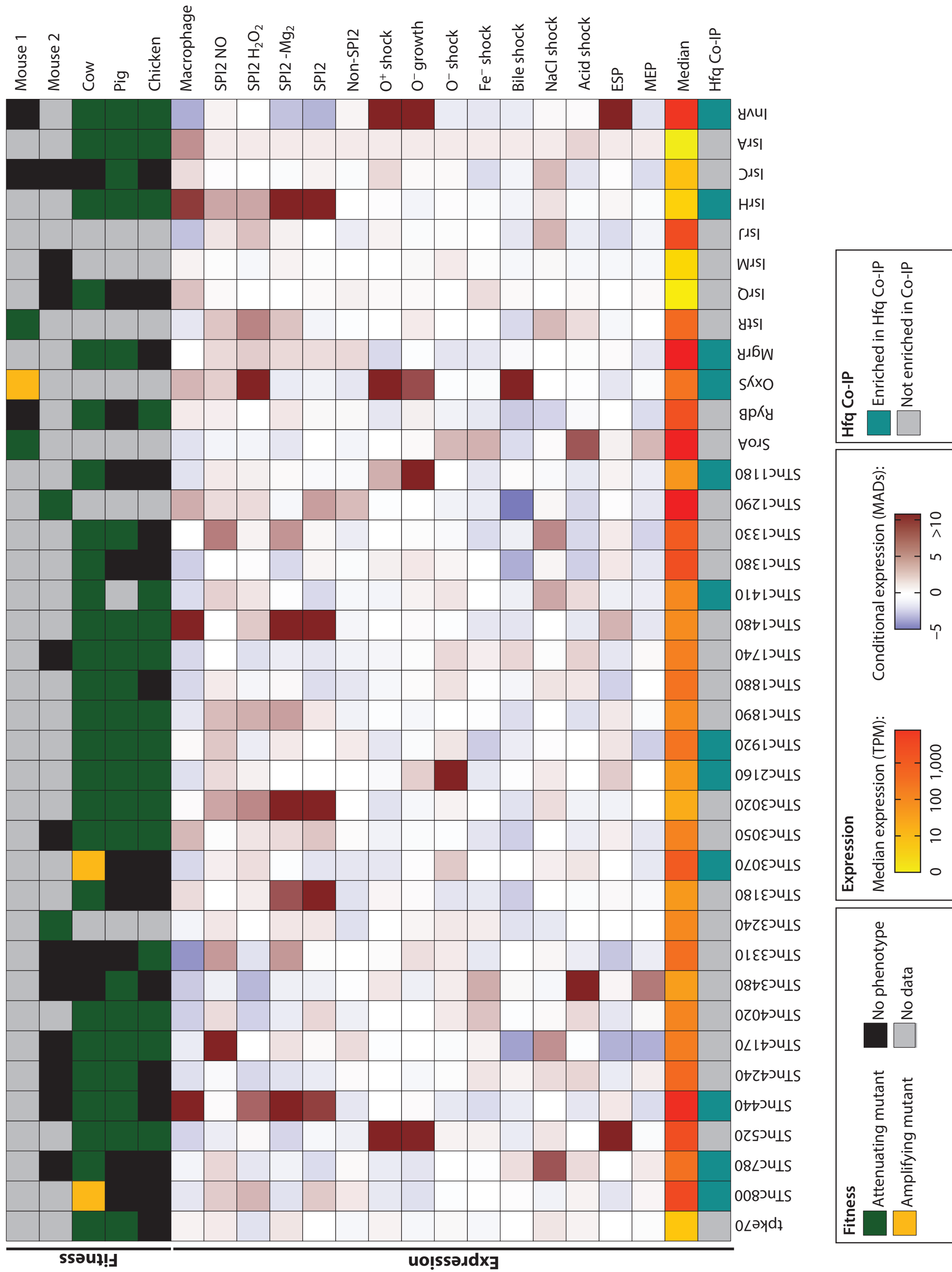
Mouse 1
Mouse 2
Cow
Pig
Chicken
Macrophage
SPI2 NO
SPI2 H_2O_2
SPI2 -Mg_2
SPI2
Non-SPI2
O^+ shock
O^- growth
O^- shock
Fe^- shock
Bile shock
NaCl shock
Acid shock
ESP
MEP
Median
Hfq Co-IP
InvR
IsrA
IsrC
IsrH
IsrJ
IsrM
IsrQ
IstR
MgrR
OxyS
RydB
SroA
STnc1180
STnc1290
STnc1330
STnc1380
STnc1410
STnc1480
STnc1740
STnc1880
STnc1890
STnc1920
STnc2160
STnc3020
STnc3050
STnc3070
STnc3180
STnc3240
STnc3310
STnc3480
STnc4020
STnc4170
STnc4240
STnc440
STnc520
STnc780
STnc800
tpke70
Fitness
Expression
Fitness
Attenuating mutant
Amplifying mutant
No phenotype
No data
Expression
Median expression (TPM):
0 10 100 1,000
Conditional expression (MADs):
−5 0 5 >10
Hfq Co-IP
Enriched in Hfq Co-IP
Not enriched in Co-IP

Figure 6

Integrating postgenomic data sets. Transcript boundaries were obtained for 280 verified and putative *Salmonella* Typhimurium sRNAs from the Hinton laboratory's SalCom (*Salmonella* Compendium) transcriptomic atlas (87), as were condition-specific expression data for a subset of 14 in vitro conditions, representative of a range of conditions *S.* Typhimurium encounters during infection, assayed in that study. These data were supplemented with RNA-seq data from intracellular *S.* Typhimurium replicating in a RAW murine macrophage-like cell line (J.C.D. Hinton, personal communication), as survival in macrophages is known to be a major determinant for the establishment of systemic infection (50). The expression level of each sRNA has been normalized to its median expression in transcripts per million (TPM) across these 14 conditions. Conditional expression has been re-expressed as median absolute deviations (MADs) from this median to provide a comparable overview of the expression dynamics of each sRNA. These sRNAs were next filtered for the presence of attenuating or amplifying mutations (fitness) in at least one of four animal models used in a transposon insertion sequencing screen (26) and in a microarray-based targeted deletion screen using a murine typhoid model (165) as well as for previous reports of virulence-related phenotypes in the literature (56, 72, 134). Additionally, sRNAs enrichment in an Hfq RIP-seq experiment (24) has been annotated. Abbreviation: Co-IP, co-immunoprecipitation.

We have also shown a brief example of how this global information can be integrated to provide guidance as to potential sRNA functions. A long-term goal for the field may be to use this sort of data to integrate sRNA networks into quantitative genome-scale models of regulatory and physiological networks (13), allowing the in silico modeling of the effects of sRNA regulation. However, for the moment, even basic questions regarding the functions of the majority of noncoding transcripts in bacteria remain unanswered, and there is much work to be done.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors would like to thank Jay Hinton, Paul Gardner, Kai Papenfort, Carsten Kröger, Rocío Canals, and members of the Vogel lab (Alex Westermann, Caroline Taouk, Yanjie Chao, and Erik Holmqvist) for helpful comments on the manuscript. L.B. is supported by a research fellowship from the Alexander von Humboldt Foundation. Work in the Vogel lab is supported by the Bavarian BioSysNet program and a BMBF grant e:Bio RNAsys.

LITERATURE CITED

1. Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G. 1997. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* 90(1):43–53
2. Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EGH, et al. 2001. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* 11(12):941–50
3. Arnvig KB, Comas I, Thomson NR, Houghton J, Boshoff HI, et al. 2011. Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of *Mycobacterium tuberculosis*. *PLOS Pathog.* 7(11):e1002342
4. Artsimovitch I, Landick R. 2002. The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* 109(2):193–203
5. Ashton PM, Nair S, Dallman T, Rubino S, Rabsch W, et al. 2014. MinIon nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat. Biotechnol.* 33:296–300

6. Bardill JP, Hammer BK. 2012. Non-coding sRNAs regulate virulence in the bacterial pathogen *Vibrio cholerae*. *RNA Biol.* 9(4):392–401
7. Barquist L, Boinett CJ, Cain AK. 2013. Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biol.* 10(7):1161–69
8. Beisel CL, Storz G. 2010. Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol. Rev.* 34(5):866–82
9. Beisel CL, Updegrove TB, Janson BJ, Storz G. 2012. Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J.* 31(8):1961–74
10. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, et al. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277(5331):1453–62
11. Bobrovskyy M, Vanderpool CK. 2013. Regulation of bacterial metabolism by small RNAs using diverse mechanisms. *Annu. Rev. Genet.* 47:209–32
12. Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, et al. 2007. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator rot by an antisense mechanism. *Genes Dev.* 21(11):1353–66
13. Bordbar A, Monk JM, King ZA, Palsson BO. 2014. Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* 15(2):107–20
14. Bossi L, Figueroa-Bossi N. 2007. A small RNA downregulates LamB maltoporin in *Salmonella*. *Mol. Microbiol.* 65(3):799–810
15. Bossi L, Schwartz A, Guillemardet B, Boudvillain M, Figueroa-Bossi N. 2012. A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes Dev.* 26(16):1864–73
16. Brantl S, Brückner R. 2014. Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol.* 11(5):443–56
17. Brouwer RWW, Kuipers OP, van Hijum SAFT. 2008. The relative value of operon predictions. *Brief. Bioinform.* 9(5):367–75
18. Brownlee GG. 1971. Sequence of 6S RNA of *E. coli*. *Nature* 229(5):147–49
19. Bryant J, Chewapreecha C, Bentley SD. 2012. Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. *Future Microbiol.* 7(11):1283–96
20. Busch A, Richter AS, Backofen R. 2008. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24(24):2849–56
21. Caldelari I, Chao Y, Romby P, Vogel J. 2013. RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb. Perspect. Med.* 3(9):a010298
22. Cambray G, Guimaraes JC, Mutalik VK, Lam C, Mai Q-A, et al. 2013. Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res.* 41(9):5139–48
23. Caron M-P, Lafontaine DA, Massé E. 2010. Small RNA–mediated regulation at the level of transcript stability. *RNA Biol.* 7(2):140–44
24. Chao Y, Papenfort K, Reinhardt R, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J.* 31(20):4005–19
25. Chao Y, Vogel J. 2010. The role of Hfq in bacterial pathogens. *Curr. Opin. Microbiol.* 13(1):24–33
26. Chaudhuri RR, Morgan E, Peters SE, Pleasance SJ, Hudson DL, et al. 2013. Comprehensive assignment of roles for *Salmonella* Typhimurium genes in intestinal colonization of food-producing animals. *PLOS Genet.* 9(4):e1003456
27. Chaudhuri RR, Peters SE, Pleasance SJ, Northen H, Willers C, et al. 2009. Comprehensive identification of *Salmonella enterica* serovar Typhimurium genes required for infection of BALB/c mice. *PLOS Pathog.* 5(7):e1000529
28. Chen Y-J, Liu P, Nielsen AAK, Brophy JAN, Clancy K, et al. 2013. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* 10(7):659–64
29. Cho B-K, Zengler K, Qiu Y, Park YS, Knight EM, et al. 2009. The transcription unit architecture of the *Escherichia coli* genome. *Nat. Biotechnol.* 27(11):1043–49
30. Conway T, Creecy JP, Maddox SM, Grissom JE, Conkle TL, et al. 2014. Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. *mBio* 5(4):e01442–14

31. Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JCD, Vogel J. 2012. Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol. Microbiol.* 84(3):428–45
32. Creecy JP, Conway T. 2014. Quantitative bacterial transcriptomics with RNA-seq. *Curr. Opin. Microbiol.* 23:133–40
33. Croucher NJ, Didelot X. 2014. The application of genomics to tracing bacterial pathogen transmission. *Curr. Opin. Microbiol.* 23:62–67
34. Croucher NJ, Thomson NR. 2010. Studying bacterial transcriptomes using RNA-seq. *Curr. Opin. Microbiol.* 13(5):619–24
35. Curtis MM, Hu Z, Klimko C, Narayanan S, Deberardinis R, Sperandio V. 2014. The gut commensal *Bacteroides thetaiotaomicron* exacerbates enteric infection through modification of the metabolic landscape. *Cell Host Microbe* 16(6):759–69
36. Darnell RB. 2010. HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip. Rev. RNA* 1(2):266–86
37. De Lay N, Schu DJ, Gottesman S. 2013. Bacterial small RNA–based negative regulation: Hfq and its accomplices. *J. Biol. Chem.* 288(12):7996–8003
38. Desnoyers G, Morissette A, Prévost K, Massé E. 2009. Small RNA–induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J.* 28(11):1551–61
39. Dimastrogiovanni D, Fröhlich KS, Bandyra KJ, Bruce HA, Hohensee S, et al. 2014. Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife* doi:10.7554/eLife.05375
40. Dugar G, Herbig A, Förstner KU, Heidrich N, Reinhardt R, et al. 2013. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLOS Genet.* 9(5):e1003495
41. Durand S, Braun F, Lioliou E, Romilly C, Helfer A-C, et al. 2015. A nitric oxide regulated small RNA controls expression of genes involved in redox homeostasis in *Bacillus subtilis*. *PLOS Genet.* 11(2):e1004957
42. Eddy SR. 2014. Computational analysis of conserved RNA secondary structure in transcriptomes and genomes. *Annu. Rev. Biophys.* 43:433–56
43. Eggenhofer F, Tafer H, Stadler PF, Hofacker IL. 2011. RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res.* 39:W149–54
44. Engreitz JM, Sirokman K, McDonel P, Shishkin AA, Surka C, et al. 2014. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell* 159(1):188–99
45. Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *PNAS* 99(11):7687–92
46. Feaga HA, Viollier PH, Keiler KC. 2014. Release of nonstop ribosomes is essential. *mBio* 5(6):e01916
47. Feasey NA, Dougan G, Kingsley RA, Heyderman RS, Gordon MA. 2012. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. *Lancet* 379(9835):2489–99
48. Feng L, Rutherford ST, Papenfort K, Bagert JD, van Kessel JC, et al. 2015. A Qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* 160(1–2):228–40
49. Ferreyra JA, Wu KJ, Hryckowian AJ, Bouley DM, Weimer BC, Sonnenburg JL. 2014. Gut microbiota–produced succinate promotes *C. difficile* infection after antibiotic treatment or motility disturbance. *Cell Host Microbe* 16(6):770–77
50. Fields PI, Swanson RV, Haidaris CG, Heffron F. 1986. Mutants of *Salmonella typhimurium* that cannot survive within the macrophage are avirulent. *PNAS* 83(14):5189–93
51. Figueroa-Bossi N, Valentini M, Malleret L, Fiorini F, Bossi L. 2009. Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev.* 23(17):2004–15
52. Friedersdorf MB, Keene JD. 2014. Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs. *Genome Biol.* 15(1):R2
53. Fröhlich KS, Papenfort K, Fekete A, Vogel J. 2013. A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J.* 32(22):2963–79
54. Fröhlich KS, Vogel J. 2009. Activation of gene expression by small RNA. *Curr. Opin. Microbiol.* 12(6):674–82

55. Gardner PP, Barquist L, Bateman A, Nawrocki EP, Weinberg Z. 2011. RNIE: genome-wide prediction of bacterial intrinsic terminators. *Nucleic Acids Res.* 39(14):5845–52
56. Gong H, Vu G-P, Bai Y, Chan E, Wu R, et al. 2011. A *Salmonella* small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. *PLOS Pathog.* 7(9):e1002120
57. Gosai SJ, Foley SW, Wang D, Silverman IM, Selamoglu N, et al. 2015. Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the *Arabidopsis* nucleus. *Mol. Cell* 57(2):376–88
58. Gottesman S. 2004. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu. Rev. Microbiol.* 58:303–28
59. Gottesman S, Storz G. 2011. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* 3(12): pii:a003798
60. Grant AJ, Restif O, McKinley TJ, Sheppard M, Maskell DJ, Mastroeni P. 2008. Modelling within-host spatiotemporal dynamics of invasive bacterial disease. *PLOS Biol.* 6(4):e74
61. Gripenland J, Netterling S, Loh E, Tiensuu T, Toledo-Arana A, Johansson J. 2010. RNAs: regulators of bacterial virulence. *Nat. Rev. Microbiol.* 8(12):857–66
62. Grosswendt S, Filipchyk A, Manzano M, Klironomos F, Schilling M, et al. 2014. Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol. Cell* 54(6):1042–54
63. Güell M, van Noort V, Yus E, Chen W-H, Leigh-Bell J, et al. 2009. Transcriptome complexity in a genome-reduced bacterium. *Science* 326(5957):1268–71
64. Güell M, Yus E, Lluch-Senar M, Serrano L. 2011. Bacterial transcriptomics: What is beyond the RNA horiz-ome? *Nat. Rev. Microbiol.* 9(9):658–69
65. Guo MS, Updegrove TB, Gogol EB, Shabalina SA, Gross CA, Storz G. 2014. MicL, a new σE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev.* 28(14):1620–34
66. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, et al. 2010. Ab initio reconstruction of cell type–specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* 28(5):503–10
67. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, et al. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141(1):129–41
68. Hammann P, Parmentier D, Cerciat M, Reimegård J, Helfer A-C, et al. 2014. A method to map changes in bacterial surface composition induced by regulatory RNAs in *Escherichia coli* and *Staphylococcus aureus*. *Biochimie* 106:175–79
69. Hammarlöf DL, Canals R, Hinton JCD. 2013. The fun of identifying gene function in bacterial pathogens; insights from *Salmonella f*unctional genomics. *Curr. Opin. Microbiol.* 16(5):643–51
70. Hansen AK, Degnan PH. 2014. Widespread expression of conserved small RNAs in small symbiont genomes. *ISME J.* 8(12):2490–502
71. He M, Miyajima F, Roberts P, Ellison L, Pickard DJ, et al. 2013. Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. *Nat. Genet.* 45(1):109–13
72. Hébrard M, Kröger C, Srikumar S, Colgan A, Händler K, Hinton JCD. 2012. sRNAs and the virulence of *Salmonella enterica* serovar Typhimurium. *RNA Biol.* 9(4):437–45
73. Helwak A, Kudla G, Dudnakova T, Tollervey D. 2013. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* 153(3):654–65
74. Hershberg R, Altuvia S, Margalit H. 2003. A survey of small RNA-encoding genes in *Escherichia coli*. *Nucleic Acids Res.* 31(7):1813–20
75. Hess WR, Berghoff BA, Wilde A, Steglich C, Klug G. 2014. Riboregulators and the role of Hfq in photosynthetic bacteria. *RNA Biol.* 11(5):413–26
76. Holmqvist E, Reimegård J, Sterk M, Grantcharova N, Römling U, Wagner EGH. 2010. Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J.* 29(11):1840–50
77. Holmqvist E, Reimegård J, Wagner EGH. 2013. Massive functional mapping of a 5′-UTR by saturation mutagenesis, phenotypic sorting and deep sequencing. *Nucleic Acids Res.* 41(12):e122

78. Holmqvist E, Unoson C, Reimegård J, Wagner EGH. 2012. A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol. Microbiol.* 84(3):414–27
79. Hussein R, Lim HN. 2011. Disruption of small RNA signaling caused by competition for Hfq. *PNAS* 108(3):1110–15
80. Ingolia NT. 2014. Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* 15(3):205–13
81. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324(5924):218–23
82. Katz Y, Wang ET, Airoldi EM, Burge CB. 2010. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods* 7(12):1009–15
83. Kim D, Hong JS-J, Qiu Y, Nagarajan H, Seo J-H, et al. 2012. Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLOS Genet.* 8(8):e1002867
84. König J, Zarnack K, Luscombe NM, Ule J. 2012. Protein-RNA interactions: new genomic technologies and perspectives. *Nat. Rev. Genet.* 13(2):77–83
85. König J, Zarnack K, Rot G, Curk T, Kayikci M, et al. 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17(7):909–15
86. Koren S, Phillippy AM. 2015. One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Curr. Opin. Microbiol.* 23:110–20
87. Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, et al. 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. *Cell Host Microbe* 14(6):683–95
88. Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, et al. 2012. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *PNAS* 109(20):E1277–86
89. Kudla G, Granneman S, Hahn D, Beggs JD, Tollervey D. 2011. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *PNAS* 108(24):10010–15
90. Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JT, Massé E. 2015. A 3 external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol. Cell.* 58(3):393–405
91. Langridge GC, Phan M-D, Turner DJ, Perkins TT, Parts L, et al. 2009. Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* 19(12):2308–16
92. Law CW, Chen Y, Shi W, Smyth GK. 2014. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15(2):R29
93. Lease RA, Cusick ME, Belfort M. 1998. Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *PNAS* 95(21):12456–61
94. Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell.* 118(1):69–82
95. Li G-W, Oh E, Weissman JS. 2012. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484(7395):538–41
96. Lindgreen S, Umu SU, Lai AS-W, Eldai H, Liu W, et al. 2014. Robust identification of non-coding RNA from transcriptomes requires phylogenetically-informed sampling. *PLOS Comput. Biol.* 10(10):e1003907
97. Livny J, Waldor MK. 2007. Identification of small RNAs in diverse bacterial species. *Curr. Opin. Microbiol.* 10(2):96–101
98. Loman NJ, Constantinidou C, Chan JZM, Halachev M, Sergeant M, et al. 2012. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat. Rev. Microbiol.* 10(9):599–606
99. Love M, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DEseq2. *Genome Biol.* 15(12):550
100. Mandin P, Gottesman S. 2009. A genetic approach for finding small RNA regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol. Microbiol.* 72(3):551–65

101. Mann B, van Opijnen T, Wang J, Obert C, Wang Y-D, et al. 2012. Control of virulence by small RNAs in *Streptococcus pneumoniae*. *PLOS Pathog.* 8(7):e1002788
102. Martin J, Zhu W, Passalacqua KD, Bergman N, Borodovsky M. 2010. *Bacillus anthracis* genome organization in light of whole transcriptome sequencing. *BMC Bioinform.* 11(Suppl. 3):S10
103. Massé E, Vanderpool CK, Gottesman S. 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J. Bacteriol.* 187(20):6962–71
104. Mastroeni P, Grant A, Restif O, Maskell D. 2009. A dynamic view of the spread and intracellular distribution of *Salmonella enterica*. *Nat. Rev. Microbiol.* 7(1):73–80
105. Mazurkiewicz P, Tang CM, Boone C, Holden DW. 2006. Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. *Nat. Rev. Genet.* 7(12):929–39
106. McAdam PR, Richardson EJ, Fitzgerald JR. 2014. High-throughput sequencing for the study of bacterial pathogen biology. *Curr. Opin. Microbiol.* 19:106–13
107. McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. *Nucleic Acids Res.* 40(10):4288–97
108. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, et al. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413(6858):852–56
109. McClure R, Balasubramanian D, Sun Y, Bobrovskyy M, Sumby P, et al. 2013. Computational analysis of bacterial RNA-seq data. *Nucleic Acids Res.* 41(14):e140
110. McHugh CA, Russell P, Guttman M. 2014. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biol.* 15(1):203
111. Mellin JR, Cossart P. 2012. The non-coding RNA world of the bacterial pathogen *Listeria monocytogenes*. *RNA Biol.* 9(4):372–78
112. Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, et al. 2009. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLOS ONE* 4(10):e7526
113. Michaux C, Verneuil N, Hartke A, Giard J-C. 2014. Physiological roles of small RNA molecules. *Microbiology* 160(Pt. 6):1007–19
114. Mika F, Hengge R. 2014. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*. *RNA Biol.* 11(5):494–507
115. Milek M, Wyler E, Landthaler M. 2012. Transcriptome-wide analysis of protein-RNA interactions using high-throughput sequencing. *Semin. Cell Dev. Biol.* 23(2):206–12
116. Mili S, Steitz JA. 2004. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 10(11):1692–94
117. Miyakoshi M, Chao Y, Vogel J. 2015. Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.* 34(11):1478–92
118. Miyakoshi M, Chao Y, Vogel J. 2015. Regulatory small RNAs from the 3′ regions of bacterial mRNAs. *Curr. Opin. Microbiol.* 24:132–39
119. Mizuno T, Chou MY, Inouye M. 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *PNAS* 81(7):1966–70
120. Modi SR, Camacho DM, Kohanski MA, Walker GC, Collins JJ. 2011. Functional characterization of bacterial sRNAs using a network biology approach. *PNAS* 108(37):15522–27
121. Møller T, Franch T, Højrup P, Keene DR, Bächinger HP, et al. 2002. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell* 9(1):23–30
122. Moon K, Gottesman S. 2011. Competition among Hfq-binding small RNAs in *Escherichia coli*. *Mol. Microbiol.* 82(6):1545–62
123. Morita T, Maki K, Aiba H. 2005. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* 19(18):2176–86
124. Moxon R, Bayliss C, Hood D. 2006. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.* 40:307–33
125. Nature. 2014. The digital toolbox. *Nature* 513(7516):6
126. Nat. Methods. 2013. Matters of significance. *Nat. Methods* 10(9):805
127. Ng W-L, Bassler BL. 2009. Bacterial quorum-sensing network architectures. *Annu. Rev. Genet.* 43:197–222

128. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, et al. 2011. Phenotypic landscape of a bacterial cell. *Cell* 144(1):143–56
129. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. 1993. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.* 12(10):3967–75
130. Oglesby-Sherrouse AG, Murphy ER. 2013. Iron-responsive bacterial small RNAs: variations on a theme. *Metallomics* 5(4):276–86
131. Oh E, Becker AH, Sandikci A, Huber D, Chaba R, et al. 2011. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 147(6):1295–308
132. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, et al. 2012. Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. *Nat. Genet.* 44(11):1215–21
133. Overgaard M, Johansen J, Møller-Jensen J, Valentin-Hansen P. 2009. Switching off small RNA regulation with trap-mRNA. *Mol. Microbiol.* 73(5):790–800
134. Padalon-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, et al. 2008. Small RNAs encoded within genetic islands of *Salmonella* Typhimurium show host-induced expression and role in virulence. *Nucleic Acids Res.* 36(6):1913–27
135. Pain A, Ott A, Amine H, Rochet T, Bouloc P, Gauthret D. 2015. An assessment of bacterial small RNA target prediction programs. *RNA Biol.* 12(5):509–13
136. Papenfort K, Bouvier M, Mika F, Sharma CM, Vogel J. 2010. Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA. *PNAS* 107(47):20435–40
137. Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JCD, Vogel J. 2006. σ^E-dependent small RNAs of salmonella respond to membrane stress by accelerating global *omp* mRNA decay. *Mol. Microbiol.* 62(6):1674–88
138. Papenfort K, Said N, Welsink T, Lucchini S, Hinton JCD, Vogel J. 2009. Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol. Microbiol.* 74(1):139–58
139. Papenfort K, Sun Y, Miyakoshi M, Vanderpool CK, Vogel J. 2013. Small RNA–mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* 153(2):426–37
140. Papenfort K, Vogel J. 2014. Small RNA functions in carbon metabolism and virulence of enteric pathogens. *Front. Cell. Infect. Microbiol.* 4:91
141. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, et al. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* 413(6858):848–52
142. Peer A, Margalit H. 2011. Accessibility and evolutionary conservation mark bacterial small-RNA target-binding regions. *J. Bacteriol.* 193(7):1690–701
143. Peer A, Margalit H. 2014. Evolutionary patterns of *Escherichia coli* small RNAs and their regulatory interactions. *RNA* 20(7):994–1003
144. Peng Y, Curtis JE, Fang X, Woodson SA. 2014. Structural model of an mRNA in complex with the bacterial chaperone Hfq. *PNAS* 111(48):17134–39
145. Pernitzsch SR, Tirier SM, Beier D, Sharma CM. 2014. A variable homopolymeric G-repeat defines small RNA–mediated posttranscriptional regulation of a chemotaxis receptor in *Helicobacter pylori*. *PNAS* 111(4):E501–10
146. Peterman N, Lavi-Itzkovitz A, Levine E. 2014. Large-scale mapping of sequence-function relations in small regulatory RNAs reveals plasticity and modularity. *Nucleic Acids Res.* 42(19):12177–88
147. Peters JM, Vangeloff AD, Landick R. 2011. Bacterial transcription terminators: the RNA 3′-end chronicles. *J. Mol. Biol.* 412(5):793–813
148. Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, Vogel J. 2007. A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol. Microbiol.* 66(5):1174–91
149. Pham TAN, Clare S, Goulding D, Arasteh JM, Stares MD, et al. 2014. Epithelial il-22ra1-mediated fucosylation promotes intestinal colonization resistance to an opportunistic pathogen. *Cell Host Microbe* 16(4):504–16
150. Pham TAN, Lawley TD. 2014. Emerging insights on intestinal dysbiosis during bacterial infections. *Curr. Opin. Microbiol.* 17:67–74

151. Pichon C, Felden B. 2005. Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *PNAS* 102(40):14249–54
152. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, et al. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341
153. Quick J, Quinlan AR, Loman NJ. 2014. A reference bacterial genome dataset generated on the MinIon^{TM} portable single-molecule nanopore sequencer. *GigaScience* 3:22
154. Rabani M, Raychowdhury R, Jovanovic M, Rooney M, Stumpo DJ, et al. 2014. High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. *Cell* 159(7):1698–710
155. Raghavan R, Groisman EA, Ochman H. 2011. Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.* 21(9):1487–97
156. Rasmussen AA, Eriksen M, Gilany K, Udesen C, Franch T, et al. 2005. Regulation of OmpA mRNA stability: the role of a small regulatory RNA in growth phase–dependent control. *Mol. Microbiol.* 58(5):1421–29
157. Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499(7457):172–77
158. Richter AS, Backofen R. 2012. Accessibility and conservation: general features of bacterial small RNA–mRNA interactions? *RNA Biol.* 9(7):954–65
159. Riley KJ, Steitz JA. 2013. The "observer effect" in genome-wide surveys of protein-RNA interactions. *Mol. Cell* 49(4):601–4
160. Rivas E, Klein RJ, Jones TA, Eddy SR. 2001. Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Curr. Biol.* 11(17):1369–73
161. Romilly C, Caldelari I, Parmentier D, Lioliou E, Romby P, Fechter P. 2012. Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus*. *RNA Biol.* 9(4):402–13
162. Sahagan BG, Dahlberg JE. 1979. A small, unstable RNA molecule of *Escherichia coli*: spot 42 RNA: I. nucleotide sequence analysis. *J. Mol. Biol.* 131(3):573–92
163. Saliba A-E, Westermann AJ, Gorski SA, Vogel J. 2014. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42(14):8845–60
164. Salzberg SL, Phillippy AM, Zimin A, Puiu D, Magoc T, et al. 2012. GAGE: a critical evaluation of genome assemblies and assembly algorithms. *Genome Res.* 22(3):557–67
165. Santiviago CA, Reynolds MM, Porwollik S, Choi S-H, Long F, et al. 2009. Analysis of pools of targeted *Salmonella* deletion mutants identifies novel genes affecting fitness during competitive infection in mice. *PLOS Pathog.* 5(7):e1000477
166. Santos RL, Zhang S, Tsolis RM, Kingsley RA, Adams LG, Bäumler AJ. 2001. Animal models of *Salmonella* infections: enteritis versus typhoid fever. *Microbes Infect.* 3(14–15):1335–44
167. Schrader JM, Zhou B, Li G-W, Lasker K, Childers WS, et al. 2014. The coding and noncoding architecture of the *Caulobacter crescentus* genome. *PLOS Genet.* 10(7):e1004463
168. Sharma CM, Darfeuille F, Plantinga TH, Vogel J. 2007. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.* 21(21):2804–17
169. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, et al. 2010. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464(7286):250–55
170. Sharma CM, Vogel J. 2009. Experimental approaches for the discovery and characterization of regulatory small RNA. *Curr. Opin. Microbiol.* 12(5):536–46
171. Sharma CM, Vogel J. 2014. Differential RNA-seq: the approach behind and the biological insight gained. *Curr. Opin. Microbiol.* 19:97–105
172. Silverman IM, Li F, Alexander A, Goff L, Trapnell C, et al. 2014. RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome. *Genome Biol.* 15(1):R3
173. Singh N, Wade JT. 2014. Identification of regulatory RNA in bacterial genomes by genome-scale mapping of transcription start sites. *Methods Mol. Biol.* 1103:1–10
174. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, et al. 2008. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLOS Genet.* 4(8):e1000163

175. Sorek R, Cossart P. 2010. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.* 11(1):9–16
176. Stevens MP, Humphrey TJ, Maskell DJ. 2009. Molecular insights into farm animal and zoonotic *Salmonella* infections. *Philos. Trans. R. Soc. Lond. B* 364(1530):2709–23
177. Storz G, Gottesman S. 2015. RNA reflections: converging on Hfq. *RNA* 21:511–512
178. Stringer AM, Currenti S, Bonocora RP, Baranowski C, Petrone BL, et al. 2014. Genome-scale analyses of *Escherichia coli* and *Salmonella enterica* AraC reveal noncanonical targets and an expanded core regulon. *J. Bacteriol.* 196(3):660–71
179. Temkin E, Adler A, Lerner A, Carmeli Y. 2014. Carbapenem-resistant Enterobacteriaceae: biology, epidemiology, and management. *Ann. N. Y. Acad. Sci.* 1323:22–42
180. Tjaden B, Goodwin SS, Opdyke JA, Guillier M, Fu DX, et al. 2006. Target prediction for small, noncoding RNAs in bacteria. *Nucleic Acids Res.* 34(9):2791–802
181. Tjaden B, Saxena RM, Stolyar S, Haynor DR, Kolker E, Rosenow C. 2002. Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Res.* 30(17):3732–38
182. Toffano-Nioche C, Nguyen AN, Kuchly C, Ott A, Gautheret D, et al. 2012. Transcriptomic profiling of the oyster pathogen *Vibrio splendidus* opens a window on the evolutionary dynamics of the small RNA repertoire in the *Vibrio* genus. *RNA* 18(12):2201–19
183. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, et al. 2009. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459(7249):950–56
184. Tomizawa J, Itoh T, Selzer G, Som T. 1981. Inhibition of colE1 RNA primer formation by a plasmid-specified small RNA. *PNAS* 78(3):1421–25
185. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with Tophat and Cufflinks. *Nat. Protoc.* 7(3):562–78
186. Tree JJ, Granneman S, McAteer SP, Tollervey D, Gally DL. 2014. Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. *Mol. Cell* 55(2):199–213
187. Tsolis RM, Xavier MN, Santos RL, Bäumler AJ. 2011. How to become a top model: impact of animal experimentation on human *Salmonella* disease research. *Infect. Immun.* 79(5):1806–14
188. Udekwu KI, Darfeuille F, Vogel J, Reimegård J, Holmqvist E, Wagner EGH. 2005. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev.* 19(19):2355–66
189. Urban JH, Vogel J. 2007. Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res.* 35(3):1018–37
190. Ursell LK, Van Treuren W, Metcalf JL, Pirrung M, Gewirtz A, Knight R. 2013. Replenishing our defensive microbes. *BioEssays* 35(9):810–17
191. Van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* 6(10):767–72
192. Van Opijnen T, Camilli A. 2012. A fine scale phenotype-genotype virulence map of a bacterial pathogen. *Genome Res.* 22(12):2541–51
193. Van Opijnen T, Camilli A. 2013. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat. Rev. Microbiol.* 11(7):435–42
194. Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jäger JG, et al. 2003. RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* 31(22):6435–43
195. Vogel J, Luisi BF. 2011. Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9(8):578–89
196. Vogel J, Sharma CM. 2005. How to find small non-coding RNAs in bacteria. *Biol. Chem.* 386(12):1219–38
197. Voigt K, Sharma CM, Mitschke J, Lambrecht SJ, Voß B, et al. 2014. Comparative transcriptomics of two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. *ISME J.* 8(10):2056–68
198. Wade JT, Grainger DC. 2014. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat. Rev. Microbiol.* 12(9):647–53
199. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, et al. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456(7221):470–76

200. Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. 2001. Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 15(13):1637–51
201. Wassarman KM, Storz G. 2000. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101(6):613–23
202. Westermann AJ, Gorski SA, Vogel J. 2012. Dual RNA-seq of pathogen and host. *Nat. Rev. Microbiol.* 10(9):618–30
203. Wright PR, Richter AS, Papenfort K, Mann M, Vogel J, et al. 2013. Comparative genomics boosts target prediction for bacterial small RNAs. *PNAS* 110(37):E3487–96
204. Wurtzel O, Sapra R, Chen F, Zhu Y, Simmons BA, Sorek R. 2010. A single-base resolution map of an archaeal transcriptome. *Genome Res.* 20(1):133–41
205. Zhang A, Altuvia S, Tiwari A, Argaman L, Hengge-Aronis R, Storz G. 1998. The OxyS regulatory RNA represses RpoS translation and binds the Hfq (hf-i) protein. *EMBO J.* 17(20):6061–68
206. Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S. 2003. Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* 50(4):1111–24