Nature Reviews Microbiology | AOP, published online 24 December 2013; doi:10.1038/nrmicro2934

PROGRESS

The excludon: a new concept in bacterial antisense RNA-mediated gene regulation

Nina Sesto, Omri Wurtzel, Cristel Archambaud, Rotem Sorek and Pascale Cossart

Abstract | In recent years, non-coding RNAs have emerged as key regulators of gene expression. Among these RNAs, the antisense RNAs (asRNAs) are particularly abundant, but in most cases the function and mechanism of action for a particular asRNA remains elusive. Here, we highlight a recently discovered paradigm termed the excludon, which defines a genomic locus encoding an unusually long asRNA that spans divergent genes or operons with related or opposing functions. Because these asRNAs can inhibit the expression of one operon while functioning as an mRNA for the adjacent operon, they act as fine-tuning regulatory switches in bacteria.

The operon model, proposed in a seminal review paper in 1961 by Jacques Monod and Francois Jacob[1], laid the foundation for understanding the principles of gene regulation in bacteria. In this model, the two researchers envisioned multigene transcriptional units for which expression is co-regulated by a repressor. Although the repressor was originally predicted to be an RNA molecule, the discovery that the Lac repressor was a protein oriented studies on bacterial gene regulation towards protein regulators. However, the interest in RNA-mediated regulation in bacteria re-emerged a decade ago, when it became clear that RNA molecules do have important regulatory roles. Today, the function of non-coding regulatory RNAs is widely studied, and RNAs are increasingly being recognized as key regulators of metabolic, physiological and pathogenic processes, as well as components of bacterial adaptive immunity, such as the recently characterized CRISPR (clustered regularly interspaced short palindromic repeats) elements[2,3].

Bacterial regulatory RNAs are generally classified into three main groups: elements that are present in the 5′ UTR of the mRNA which they regulate (for example, riboswitches, thermosensors and pH sensors); *trans*-encoded small RNAs (sRNAs), which are defined as regulators of one or several target genes located elsewhere on the chromosome; and *cis*-encoded antisense RNAs (asRNAs), which overlap and are complementary to their target genes encoded on the opposite DNA strand of the same genomic locus. The regulatory function of non-coding RNAs is often tightly associated with the activity of RNases (enzymes that cleave RNA and are involved in RNA processing, degradation and quality control). Following an interaction with its target transcript, a non-coding RNA can recruit a particular RNase and promote specific cleavage and degradation of the transcript. Alternatively, a non-coding RNA can prevent transcript recognition by RNases and protect the transcript from degradation, thereby increasing the stability of the target RNA in the cell[4]. The numerous studies on *trans*-encoded sRNAs and regulatory 5′ UTR elements have been recently reviewed[5–7]. The less studied, asRNA-mediated mechanisms of gene regulation have also been reviewed[8,9]; however, this is a fast-moving field, and a novel mechanism of asRNA-mediated regulation has recently emerged from transcriptomic studies in *Listeria* spp.[10,11]. In this Progress article, we present our current knowledge of what we have termed the excludon paradigm of asRNA-mediated gene regulation[11] and put this paradigm in the context of other, better characterized asRNA-mediated regulatory mechanisms. The excludon concept describes unusually long asRNAs that inhibit the expression of one group of genes while enhancing the expression of a second group of genes. Thus, single transcripts have the ability to control divergent operons that often have opposing functions.

asRNAs in microbial transcriptomes

asRNAs are encoded on one strand of the DNA and overlap a gene that is encoded on the opposite strand. Therefore, these *cis*-encoded asRNAs have perfect complementarity to the sense transcript from the opposite DNA strand. The regulatory role of asRNAs was first reported more than 30 years ago, in the case of plasmid- and transposon-encoded asRNAs in *Escherichia coli*, when the asRNAs RNAI and CopA were found to negatively regulate plasmid copy number[12,13]; a few years later, RNA-OUT was found to regulate the transposition of the transposon Tn*10* by repressing synthesis of the transposase[14]. Following this, chromosomally encoded asRNAs were only rarely identified and considered to be exceptions rather than the rule. Indeed, by 2007, only about ten bacterial asRNAs had been described[15]. The number of reported asRNAs has recently exploded with the emergence of high-throughput RNA-seq (RNA sequencing) and tiling array studies, which have revealed an unexpected abundance of hundreds of asRNA transcripts in microbial transcriptomes. These asRNAs were found in a wide range of bacteria and archaea, such as *Mycoplasma pneumoniae*[16], *Sulfolobus solfataricus*[17], *Helicobacter pylori*[18], *Synechocystis* sp. PCC 6803 (REF. 19), *Listeria* spp.[10,20,21], *Bacillus subtilis*[22], *Salmonella enterica* subsp. *enterica* serovar Typhimurium[23], *Agrobacterium tumefaciens*[24], *Pseudomonas aeruginosa*[25] and others. The asRNAs identified in these species overlap 1–25% of protein-coding genes and up to 46% in *H. pylori*[18].

Our current knowledge of asRNAs is largely descriptive and relates to their size and genomic organization, but in most cases not their function. Generally, asRNAs exist as autonomous transcripts with promoters located on the DNA strand

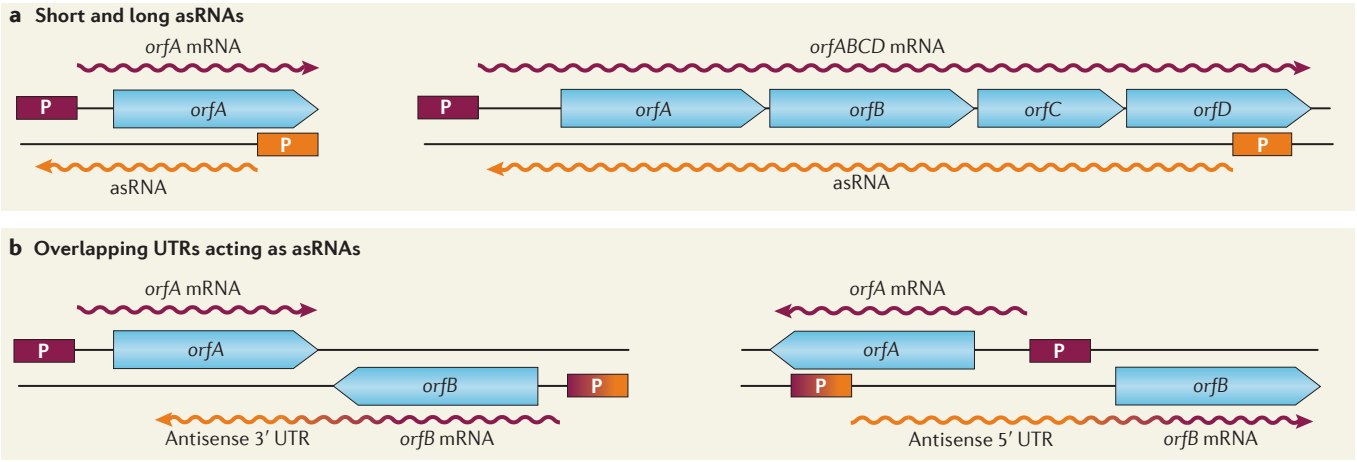

Figure 1 | **Various types of bacterial antisense RNAs.** Antisense RNAs (asRNAs) exist as autonomous transcripts of various sizes. Promoter (P) regions are indicated, blue arrows represent annotated protein-coding genes (ORFs), and red and orange arrows depict mRNA and asRNA, respectively. **a** | Short asRNAs overlap one sense ORF, whereas long asRNAs overlap several sense ORFs. **b** | An asRNA to a particular gene can also result from a long 3′ UTR (left) or 5′ UTR (right) of the mRNA transcribed from a neighbouring gene.

opposite to the ORF of the sense transcript (FIG. 1a). These RNAs vary greatly in size, ranging from short transcripts, such as the 77-nucleotide (nt) SymR asRNA that overlaps the SOS-induced gene encoding the endoribonuclease toxin SymE in *E. coli*[26] (discussed later), to long transcripts that can reach several kilobases in size. For example, in *Listeria monocytogenes*, the 2 kb asRNA Anti2095 overlaps four genes (*lmo2095*, *lmo2096*, *lmo2097* and *lmo2098*, encoding a putative phosphofructokinase and three components of the galactitol-specific phosphotransferase system, respectively)[10], and in *Prochlorococcus* sp. MED4, a 7 kb asRNA overlaps 14 genes, including several ribosomal protein genes, belonging to two adjacent operons spanning ORFs *PMM1533* to *PMM1558* (REF. 27).

Moreover, asRNAs can originate from long 5′ or 3′ UTRs of mRNAs that overlap one or several genes encoded on the opposite strand (FIG. 1b). In cyanobacteria of the genus *Anabaena*, the long 3′ UTR of the *alr1690* gene (encoding a putative cell wall-binding protein) extends and entirely overlaps the adjacent *all1691* gene (encoding the ferric uptake regulator FurA) on the opposite strand[28]. Similar examples of long UTRs acting as asRNAs have also been described in *L. monocytogenes*[10,11] and *B. subtilis*[29].

In addition to regulating protein-coding genes, asRNAs can overlap and affect the expression of non-coding sRNAs, as exemplified by the asRNA SraC (also known as RyeA) in *E. coli* and *S.* Typhimurium, which is encoded opposite SdsR (also known as RyeB), an sRNA that regulates the expression of the porin OmpD in *S.* Typhimurium[30,31].

Despite the growing collection of asRNAs reported in bacteria and archaea, the exact functions of the vast majority of these transcripts are largely unknown[32], and it is still a matter of debate whether all asRNAs function as regulatory RNAs or whether some of them represent transcriptional noise or experimental artefacts[8,33]. Below, we summarize our current knowledge of those asRNAs with defined effects and mechanisms of action. We then discuss the excludon concept and describe two examples for which experiments support their role as molecular switches.

Mechanisms of asRNA-mediated regulation

An asRNA can affect the expression of its complementary gene at the levels of transcription, mRNA stability or translation. Well-studied examples are described below and depicted in FIG. 2.

Regulation of transcription. asRNAs regulate transcription of the complementary mRNA by two main mechanisms: transcription interference and transcription attenuation.

Transcription interference involves physical and direct suppression of transcription by another transcription event. This can occur when two promoters are oriented face to face (so-called convergent promoters), such that their transcripts partially overlap and transcription events converge. Hence, when an antisense promoter is positioned face to face relative to the sense promoter, transcription in the antisense direction interferes with transcription of the sense gene. In this scenario, the transcription process itself has a regulatory role, and the asRNA is merely a by-product. This interference mechanism is exemplified by the convergent lytic promoter (*pR*) and lysogenic promoter (*pL*) of coliphage 186 (FIG. 2a). These promoters generate transcripts that overlap by 62 nt. The transcriptional activity of *pR* is intrinsically stronger than the activity of the weaker *pL* promoter: RNA polymerase binds to *pL* efficiently, but is slow to switch from an open complex to an elongating complex. By altering the arrangement of the two promoters or terminating transcription from *pR* before the RNA polymerase reached *pL*, Callen *et al.* demonstrated that the transcription complex originating from the stronger *pR* promoter elongates over *pL* and displaces the slower *pL*-bound polymerase complex, resulting in a 5.6-fold reduction in activity at *pL*. The authors termed this phenomenon the sitting-duck mechanism of transcription interference[34]. Another mechanism of interference involves the head-on collision of the two converging elongation complexes, leading to premature termination of one or both transcription processes. This mechanism was suggested to occur in *Clostridium acetobutylicum*, in which expression of the *ubiG–mccBA* operon (encoding products involved in the conversion of methionine to cysteine) is suppressed by the transcription of a long, riboswitch-regulated asRNA[35]. Finally, promoter occlusion occurs when an RNA polymerase from a stronger promoter blocks the formation of a transcription initiation complex at a weaker convergent or tandem promoter[36].

asRNA-mediated transcription attenuation occurs when a terminator structure

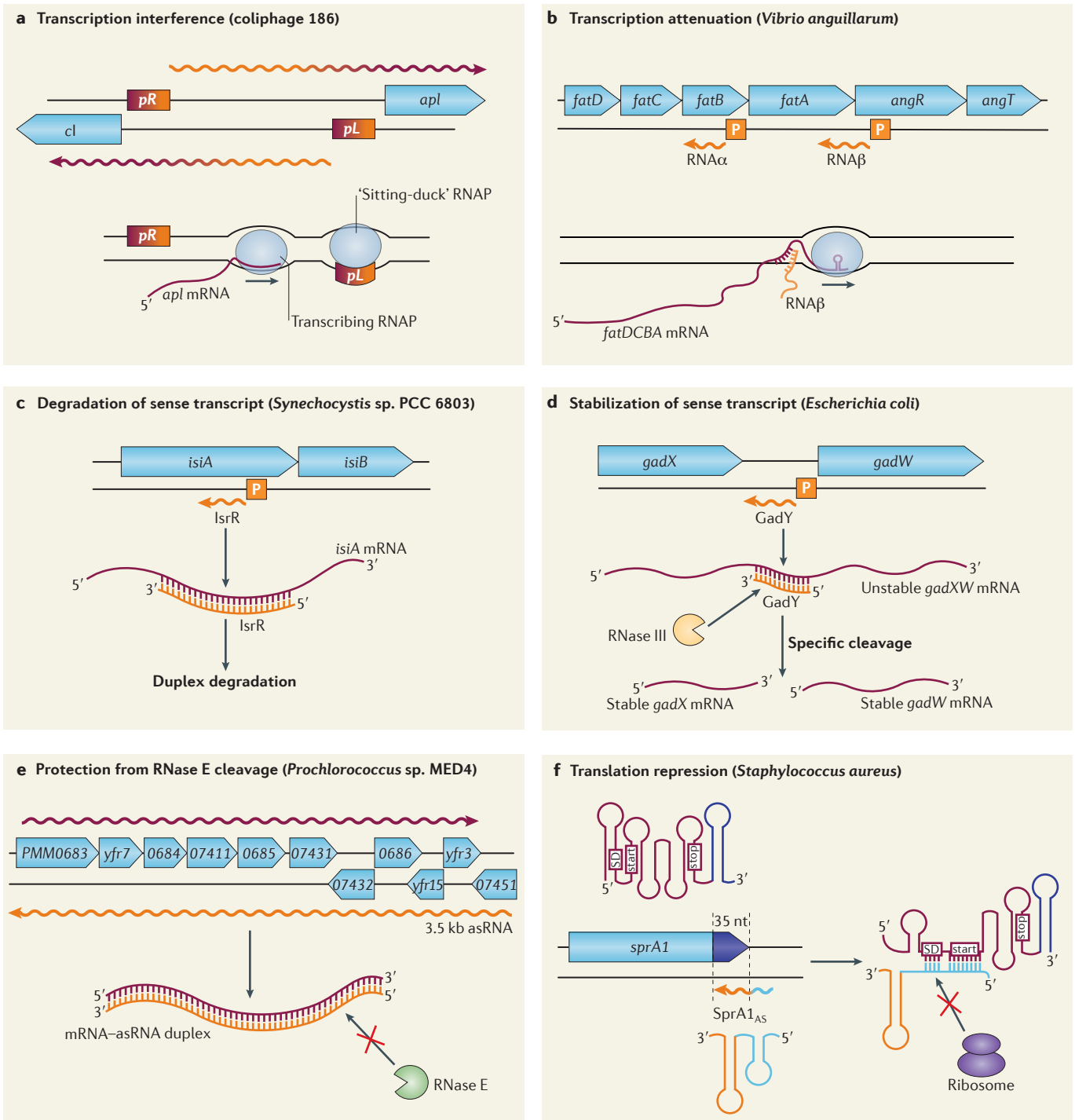

Figure 2 | **Bacterial antisense RNAs have many modes of regulating gene expression.** **a** | Transcription interference in coliphage 186 arises when the elongating RNA polymerase complex that was initiated at the strong lytic promoter (*pR*) displaces the slow 'sitting-duck' RNA polymerase (RNAP) of the weak lysogenic promoter (*pL*), resulting in downregulation of the *cI* gene. **b** | Binding of the asRNA RNAβ to the growing *fatDCBA* mRNA in *Vibrio anguillarum* induces transcription termination after the *fatA* gene, resulting in differential expression of the *fatDCBA–angRT* operon, with high expression of *fatDCBA* and low expression of *angRT*. RNAα is an asRNA that is expressed from the same locus but represses *fatA* and *fatB* expression. **c** | Negative regulation of the *Synechocystis* sp. PCC 6803 gene *isiA* (encoding iron stress-induced chlorophyll-binding protein) occurs when the asRNA IsrR is expressed and forms an RNA duplex with the *isiA* mRNA; this duplex is subsequently degraded by an unknown mechanism. **d** | Binding of the *Escherichia coli* asRNA GadY to the unstable polycistronic *gadXW* mRNA induces a specific, RNase III-mediated cleavage within the *gadXW* transcript, leading to the generation of stable monocistronic *gadX* and *gadW* transcripts. **e** | In *Prochlorococcus* sp. MED4, the interaction of a 3.5 kb asRNA with its complementary polycistronic operon masks RNase E recognition sites in the sense transcript and thus stabilizes it. **f** | Translation of the mRNA encoding the toxin SprA1 is inhibited by binding of the asRNA $SprA1_{AS}$ to the Shine–Dalgarno (SD) sequence and the start codon in the *sprA1* mRNA. Although the *sprA1* mRNA and $SprA1_{AS}$ overlap by 35 nucleotides (nt), their interaction is mediated by non-complementary regions; hence, this asRNA operates in *trans*.

is formed in the target mRNA following an interaction with an asRNA. This is exemplified by the *Vibrio anguillarum* polycistronic mRNA *fatDCBA–angRT*, which encodes ferric-siderophore transport proteins in the *fatDCBA* region and siderophore biosynthesis proteins in the *angRT* region (FIG. 2b). Although the *fatDCBA–angRT* genes are transcribed as a single operon, the expression level of the shorter *fatDCBA* mRNA is 17-fold higher than the expression level of the full-length *fatDCBA–angRT* mRNA. This differential expression is mediated by RNAβ, a 427 nt asRNA that spans the intergenic region between *fatA* and *angR* and overlaps the 3′ and 5′ UTRs of these genes, respectively. Binding of RNAβ to the growing polycistronic mRNA induces transcription termination after the *fat* genes, resulting in attenuated transcription of the distal, *angRT* region of the operon[37].

Transcription interference and transcription attenuation can also occur simultaneously. For example, in the pathogenic bacterium *Shigella flexneri*, transcription of the asRNA RnaG interferes with transcription from the weaker convergent promoter and reduces the expression of the virulence factor IcsA. In addition to this interference mechanism, the RnaG transcript can interact with the *icsA* mRNA and induce the formation of a terminator structure within its 5′ end to attenuate transcription[38].

Degradation of sense transcripts. asRNAs can induce the degradation of their complementary sense transcripts. This has been reported to occur in the cyanobacterium *Synechocystis* sp. PCC 6803, in which photosynthesis depends on available iron in the ocean. *Synechocystis* spp. respond to iron deficiency by expressing IsiA (iron stress-induced chlorophyll-binding protein), which forms a giant ring structure around photosystem I and enhances light absorption by this photosystem to compensate for the reduction in the number of photosystems that occurs under iron limitation. IsiA expression is post-transcriptionally regulated by IsrR, an antisense RNA complementary to the *isiA* mRNA (FIG. 2c). When both the *isiA* and IsrR RNAs are produced, they interact to form a double-stranded RNA duplex, which makes them a target for selective degradation[39]. Although the mechanism responsible for the rapid decay of *isiA*–IsrR RNA duplexes is unknown, other studies have demonstrated that asRNA-induced degradation can involve several RNases. For example, in *S.* Typhimurium, the asRNA AmgR destabilizes the *mgtCBR* operon in an RNase E-dependent manner and thus negatively regulates the protein levels of both MgtC (a membrane protein that is important for growth under low-Mg^{2+} conditions and also acts as a virulence factor) and MgtB (a Mg^{2+} transporter)[40]. A deep-sequencing study of *Staphylococcus aureus* reported a collection of short RNAs generated by RNase III-dependent digestion of overlapping mRNA–asRNA transcripts across the bacterial genome, suggesting that asRNAs have a general role in mediating RNA degradation in *S. aureus* and possibly other Gram-positive bacteria[41]. In the same bacterium, a co-immunoprecipitation study that used a catalytically inactive form of RNase III revealed that this enzyme associates with asRNAs that are complementary to 44% of the annotated mRNAs, and for at least one mRNA–asRNA pair, it was shown that RNase III-dependent degradation contributes to controlling the mRNA level[42].

Cleavage and stabilization of sense transcripts. asRNA–mRNA interactions can also trigger specific cleavage of the sense transcript, as exemplified by GadY, an *E. coli* asRNA that regulates the expression of the *gadXW* operon, which encodes two transcriptional regulators involved in the acid stress response (FIG. 2d). GadY is transcribed from a locus in the intergenic region between the *gadX* and *gadW* ORFs, on the opposite DNA strand. When GadY is expressed, it forms a GadY–*gadXW* duplex, which promotes a specific cleavage event involving RNase III. This cleavage results in separation of the polycistronic *gadXW* mRNA into two independent transcripts. These monocistronic transcripts accumulate at higher levels than the full-length mRNA, indicating that GadY positively regulates the expression of these two transcripts. The exact mechanism of GadY-mediated positive regulation of *gadXW* is not yet understood. Binding of GadY to the full-length transcript, or the cleavage event that this stimulates, might interfere with the recognition of instability determinants in the *gadX* and *gadW* transcripts, such as recognition sites for RNases. Alternatively, the activity of GadY might alter the secondary structure of the transcripts so that they are less susceptible to degradation[43–45].

asRNA-mediated protection of the sense transcript from degradation by RNases was reported in the cyanobacterium *Prochlorococcus* sp. MED4. Transcriptome analysis of a phage-infected *Prochlorococcus* sp. MED4 strain revealed a global decrease in the expression of most cellular mRNAs compared with their expression in an uninfected strain. However, there was a concurrent increase in the expression of a specific set of transcripts that mapped to distinct chromosomal regions. Interestingly, some of

Glossary

CRISPR
(Clustered regularly interspaced short palindromic repeats). DNA sequences comprising multiple short direct repeats. These sequences, together with CRISPR-associated (Cas) proteins, constitute an adaptive immune system that is encoded by many bacteria and most archaea, and targets invading bacteriophages and conjugative plasmids.

Gnotobiotic mice
Mice that are born in germ-free conditions and are then experimentally colonized with a defined microbiota or bacterium. Derived from the Greek roots *gnostos*, meaning known, and *bios*, meaning life.

Phosphotransferase system
A multiprotein translocation system that is only present in bacteria and catalyses the transport and phosphorylation of numerous monosaccharides, disaccharides, amino sugars, polyols and other sugar derivatives.

Photosystem
A protein complex located in the thylakoid membranes of plants, algae and cyanobacteria and in the cytoplasmic membrane of photosynthetic bacteria. Photosystem complexes carry out the primary photosynthesis reactions: the absorption of light, the transfer of electrons and the production of energy.

Pleiotropic *trans*-acting asRNAs
Antisense RNAs (asRNAs) that affect the expression of multiple genes located at different loci in the chromosome.

Riboswitches
RNA structures that are found in the 5′ UTRs of mRNAs and directly bind metabolites, metals or tRNAs to regulate the expression of the downstream sequences (which usually encode proteins involved in the synthesis or catabolism of the regulatory metabolite).

RNA-seq
(RNA sequencing). High-throughput sequencing of the cDNA obtained by the reverse transcription of an RNA pool. The sequences obtained can be used to generate a quantitative genome-wide transcriptome map of an organism.

Tiling array
A DNA microarray chip for which the hybridization probes are designed to overlap each other and to cover the whole genome (tiled probes), as opposed to a gene expression array, for which probes are usually designed to represent only the coding sequences of the genome and there are only a few probes per known gene.

Transposase
An enzyme that catalyses the excision of a transposon.

Transposon
A mobile DNA element that can excise and insert into a different locus in the genome.

these highly expressed regions encoded long asRNAs of 3.5 kb and 7 kb. These asRNAs were shown to protect their complementary mRNAs from RNase E-mediated cleavage, as the formation of an asRNA–mRNA duplex masked the single-stranded RNase E recognition sites[27] (FIG. 2e). A similar example has been described in *Synechocystis* sp. PCC 6803 involving two asRNAs, PsbA2R and PsbA3R, that overlap the 5′ UTRs of the *psbA2* and *psbA3* genes, respectively, which encode reaction centre proteins of photosystem II. Suppression of the asRNA PsbA2R results in a loss of stability and a decrease in the abundance of the target *psbA2* mRNA, consequently lowering the maximal photosystem activity. PsbA2R overlaps the RNase E recognition site (AU-box) in the *psbA2* 5′ UTR, but does not overlap the Shine–Dalgarno (SD) sequence, which is also an RNase E cleavage site. Initial *in vitro* assays showed no PsbA2R-mediated protection of *psbA2* mRNA. However, when the assay was carried out with a longer, artificial form of PsbA2R that overlapped both the AU-box and the SD sequence, the *psbA2* mRNA was protected from RNase E-mediated cleavage. On the basis of these results, the authors suggested that *psbA2* mRNA stability *in vivo* is determined by both PsbA2R and ribosomes, which protect the AU-box and the SD sequence, respectively, from RNase E-mediated cleavage[46].

Regulation of translation. Analogous to *trans*-encoded regulatory sRNAs, which often bind the SD sequence of their target mRNAs, asRNAs can downregulate translation. This mechanism has been described for the *E. coli* asRNA SymR, which is encoded opposite to the gene *symE*[26]. This gene encodes an SOS-induced toxin-like endoribonuclease that causes RNA degradation and decreased protein synthesis when overexpressed, resulting in reduced bacterial growth. SymE synthesis is tightly regulated by the antitoxin asRNA SymR, which overlaps the 5′ UTR of the *symE* mRNA, including the SD sequence and the AUG start codon. The interaction between SymR and the *symE* mRNA renders the SD sequence inaccessible to the ribosome and blocks production of SymE. Interestingly, in the absence of SymR, levels of the *symE* transcript are only mildly increased, whereas a much greater difference is observed in the levels of SymE protein. This would suggest that SymR has only a moderate effect on transcript stability and that its primary function is translation inhibition.

***asRNAs and regulation in* trans.** Although most of the asRNAs that have been studied so far are *cis* regulators, several lines of evidence indicate that asRNAs might also act on *trans*-encoded targets. This suggestion is supported by a recent study in *S. aureus* that describes the regulation of a toxin–antitoxin module. SprA1 is a toxin that functions as both a virulence factor and a growth inhibitor of *S. aureus*. Thus, expression of the toxin alone is detrimental to *S. aureus* survival. However, toxin synthesis is repressed by the co-expression of the asRNA $SprA1_{AS}$, which overlaps the 3′ end of the *sprA1* mRNA by 35 nt (FIG. 2f). Surprisingly, repression of *sprA1* is not mediated by complementary base-pairing between the two overlapping RNAs, but rather by the interaction of $SprA1_{AS}$ with the 5′ end of the *sprA1* mRNA in *trans*, leading to sequestration of the *sprA1* mRNA SD sequence and inhibition of protein synthesis[47]. This example demonstrates that a functional interaction between an asRNA and its target is not necessarily determined by the region of complementarity; therefore, regulatory asRNAs have the potential to act on multiple targets in *trans*.

The existence of such pleiotropic *trans*-acting asRNAs is further supported by a study in *L. monocytogenes*, in which the asRNA RliE partially overlaps, in *cis*, the *comC* gene, which encodes a putative competence regulator. As first predicted and then demonstrated by band-shift assays, RliE can interact in *trans* with the *comEA–comEB–comEC*, *comFA–comFC* and *lmo0945* mRNAs, which also encode putative factors of the competence machinery. For all these targets, it is the same sequence of RliE that is predicted to mediate the interaction[20]. A similar case was reported in *Mycobacterium tuberculosis*, in which the two asRNAs, ASdes and ASpks, have complementarity to multiple *M. tuberculosis* genes, suggesting that they have the ability to act in *trans*[48]. Another well-studied example is the ArcZ asRNA in *E. coli*; this asRNA negatively regulates the expression of *arcB*, which is located on the opposite strand and encodes the histidine kinase of the aerobic respiration control two-component system (ArcA–ArcB). In addition to being an asRNA, ArcZ acts as an sRNA and interacts in *trans* with the *rpoS* mRNA to activate translation of RpoS, a σ-factor subunit of RNA polymerase and the master regulator of the general stress response[49]. ArcZ is conserved in *S.* Typhimurium, in which it regulates many *trans*-encoded mRNA targets[50].

The excludon

Recent transcriptome-wide studies in *L. monocytogenes* revealed a collection of unusually long asRNAs that share a common feature: these long asRNAs (lasRNAs) not only encompass several ORFs but also extend over adjacent genes that have divergent orientations, which defines them as a group of transcripts exhibiting a unique regulatory pattern[11] (FIG. 3a). The divergently encoded genes have either related or, often, opposing functions. Transcription of these lasRNAs is initiated from a promoter located opposite a specific ORF, generating an antisense product that overlaps that ORF. However, transcription does not terminate at the end of the overlapped gene, but further extends through the neighbouring divergently oriented gene. This means that the lasRNA has a dual function, serving as an antisense regulator of one ORF or operon while simultaneously encompassing the 5′ UTR and coding sequence of the adjacent divergent ORF or operon. In other words, the antisense component of the lasRNA negatively affects the expression of one gene, whereas the distal part of the lasRNA functions as an mRNA and positively contributes to the expression of the adjacent gene. Hence, a single transcript governs the mutually exclusive expression of two adjacent genes that have opposing functions, thereby ensuring that the expression of one results in the repression of the other. We have recently coined the term excludon to describe this unique locus organization harbouring a lasRNA and two operons of mutually exclusive or related functions[11].

The first excludon reported in *L. monocytogenes* controls flagellum biosynthesis at the motility gene repressor (*mogR*; *lmo0674*)/*lmo0675–lmo0676–lmo0677* locus, in which the *mogR* and *lmo0675–lmo0676–lmo0677* genes are divergently arranged[10] (FIG. 3b). MogR is a transcriptional repressor of flagellum and motility genes in *Listeria* spp., and the *mogR* gene is expressed from a promoter located 45 nt upstream of its start codon. *lmo0675* encodes a protein with an unknown function, and *lmo0676* and *lmo0677* encode FliP and FliQ, respectively. Together with FliR (encoded by the adjacent gene, *lmo0678*), FliP and FliQ form the flagellum export apparatus and are transcriptionally regulated by MogR (which binds the *fli* operon promoter). In addition, transcription from the third promoter in this locus results in an excludon transcript, Anti0677, in which the proximal region is antisense to *lmo0675–lmo0676–lmo0677* and the distal region contains the coding sequence for MogR. As first shown by tiling arrays, Anti0677 is regulated by RNA polymerase σ-factor σ^B (a stress-activated

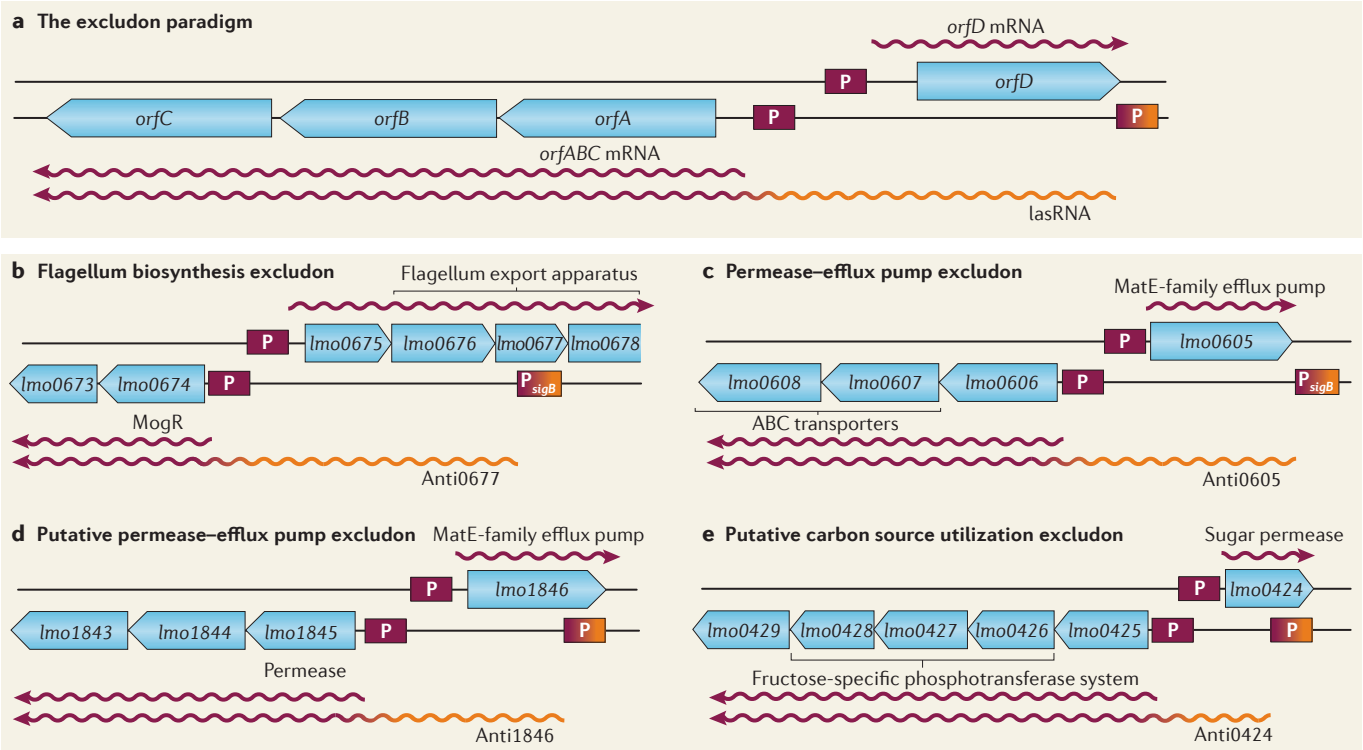

Figure 3 | **Schematic representations of excludons.** **a** | The excludon paradigm. A general representation of an excludon locus, which consists of divergently oriented genes overlapped by a long antisense RNA (lasRNA), is shown. The overlapping lasRNA can act as a negative regulator for genes encoded on the opposite strand, but it can also be used as an mRNA for the genes encoded on the same strand. **b**–**e** | Examples of excludons in *Listeria monocytogenes*. **b** | The flagellum biosynthesis excludon is regulated by the lasRNA Anti0677, which downregulates the expression of flagellum export apparatus genes (*lmo0675*–*lmo0676*–*lmo0677*) and concurrently drives expression of the gene encoding motility gene repressor (MogR), as the distal region of Anti0677 contains the coding sequence for this flagellum biosynthesis repressor. **c** | The permease–efflux pump excludon at the *lmo0605*/*lmo0606*–*lmo0607*–*lmo0608* locus in *L. monocytogenes* is regulated by the lasRNA Anti0605, which overlaps and possibly negatively affects the expression of a gene encoding a MatE-family efflux pump (*lmo0605*) while acting as a second mRNA for the divergently oriented operon encoding two ABC transporters (*lmo0607* and *lmo0608*). **d** | A second putative permease–efflux pump excludon in *L. monocytogenes* involves the lasRNA Anti1846, which overlaps the MatE-family efflux pump gene *lmo1846* and acts as an mRNA for the permease encoded by *lmo1845*. **e** | *L. monocytogenes* also contains a possible carbon source utilization excludon involving the lasRNA Anti0424. This transcript overlaps a divergent sugar permease gene (*lmo0424*; predicted to be involved in glucose uptake) and then extends into an operon (encoded on the same strand as the lasRNA) containing genes encoding components of the fructose-specific phosphotransferase system (*lmo0426*–*lmo0427*–*lmo0428*).

transcriptional regulator), and the σ^B-responsive promoter is located within the *lmo0677* ORF. In an *L. monocytogenes* strain carrying a point mutation that abolishes the activity of the σ^B-responsive Anti0677 promoter without affecting the *lmo0677* ORF, the level of MogR expression is lower than in wild-type cells, whereas the expression of FliP is higher. Thus, when Anti0677 is transcribed, it negatively regulates the expression of the sense transcript, and thereby inhibits the synthesis of the lmo0675, FliP and FliQ flagellum export apparatus while concurrently driving the expression of MogR independently of the promoter located immediately upstream of *mogR*. Two independent mechanisms — inhibition mediated by the antisense component of Anti0677 and repression mediated by increased expression of MogR — ensure that flagellum production is switched off. The opposing regulation of *mogR* and *fliP* has also been observed in an *L. monocytogenes* mutant strain with a deletion of *sigB* (the gene encoding σ^B); this strain does not express anti0677 and, interestingly, displays increased motility[10].

A recent comparative transcriptomic study of *L. monocytogenes* and *Listeria innocua*[11] detected at least 13 genomic loci that exhibit the same gene topology and lasRNA organization as the *mogR*/*lmo0675*–*lmo0676*–*lmo0677* locus, and this finding led to the definition of the excludon paradigm. Expression of three such lasRNAs (Anti0605, Anti1846 and Anti0424) was examined by northern blot, and the excludon *lmo0605*/*lmo0606*–*lmo0607*–*lmo0608* was studied in detail. The three excludons described below are conserved across, and in some cases beyond, the genus *Listeria*, thus strengthening their functional relevance.

Permease–efflux pump excludons. The genomic locus *lmo0605*/*lmo0606*–*lmo0607*–*lmo0608* contains the *lmo0605* gene, which encodes a multidrug efflux pump of the MatE family, and the divergently oriented *lmo0606*–*lmo0607*–*lmo0608* operon, which encodes a transcriptional regulator (*lmo0606*) and two putative ABC-family transport systems (*lmo0607* and *lmo0608*), each containing a permease and an ATP-binding domain (FIG. 3c). This locus encodes the 5.8 kb lasRNA Anti0605, for which the proximal region overlaps and is complementary to *lmo0605* and the distal region acts as an mRNA of the *lmo0606*–*lmo0607*–*lmo0608* operon. Similar to Anti0677, expression of Anti0605 is regulated by σ^B, and the consensus σ^B recognition sequence is located within the Anti0605 promoter region. In an *L. monocytogenes* mutant strain with a *sigB* deletion, in which Anti0605 is

not transcribed, expression of *lmo0605* is upregulated, whereas expression of the divergently encoded *lmo0606–lmo0607–lmo0608* operon is downregulated. It is not known whether the *sigB* deletion could have an indirect effect on the expression of genes in the excludon locus independently of the loss of Anti0605 expression. However, the analogy with Anti0677, for which the direct effect on genes in the excludon locus has been demonstrated, would suggest that transcription of Anti0605 leads to a direct block in the expression of the MatE-family multidrug efflux pump and a simultaneous enhancement in the expression of the ATP-driven permeases. As efflux pumps usually export molecules from the inside to the outside of the cell, whereas permeases drive the flow of molecules in the opposite direction, it is tempting to speculate that Anti0605 acts as a switch to control an import-versus-export decision for specific molecules in *Listeria* spp. A tiling array analysis was carried out using RNA extracted from bacteria grown in conditions that mimic the infectious process of *Listeria* spp., and interestingly, an inverse correlation in the expression of these two transport systems was observed: in human blood, the MatE-family efflux pump was expressed, whereas the ATP-driven permeases were repressed; conversely, in the intestinal lumen of gnotobiotic mice, a reversal in the expression pattern was observed[10].

A second putative permease–efflux pump excludon was found at the *lmo1846/lmo1845–lmo1844–lmo1843* locus in *Listeria* spp., in which the lasRNA Anti1846 is encoded on the opposite strand to the *lmo1846* gene, which encodes another MatE-family efflux pump. This lasRNA also extends into the coding sequence of a xanthine-uracil permease (encoded by *lmo1845*) (FIG. 3d). This organization suggests that Anti1846 induces expression of the permease and simultaneously represses expression of the efflux pump, again suggesting a simple mechanism for an import-versus-export cellular decision. Although the efflux pumps and the permeases in the two excludons (*lmo0605/lmo0606–lmo0607–lmo0608* and *lmo1846/lmo1845–lmo1844–lmo1843*) are not paralogous, the similar genomic organization indicates that the excludon might be a common general mechanism for controlling import and export decisions in bacteria.

A possible carbon source utilization excludon. The genomic locus *lmo0424/lmo0425–lmo0426–lmo0427–lmo0428–lmo0429* encodes carbon utilization genes; *lmo0424* is predicted to encode a sugar permease (putatively involved in glucose uptake), and the neighbouring operon encodes components of the fructose-specific phosphotransferase system (lmo0426, lmo0427 and lmo0428) (FIG. 3e). The lasRNA Anti0424 is a 6.5 kb transcript that is encoded opposite the *lmo0424* gene and continues into the adjacent operon, possibly facilitating the switch between the utilization of the two carbon sources. Selective carbon source usage is a common regulatory phenomenon that occurs in bacteria and is mediated by carbon catabolite repression, wherein the presence of one preferred carbon source induces the repression of genes involved in the consumption of secondary carbon sources[51]. The putative *lmo0424/lmo0425–lmo0426–lmo0427–lmo0428–lmo0429* excludon might represent an asRNA-based mechanism for selective carbon usage, but this hypothesis requires further experimental investigation.

Potential mechanism of excludon action. Experiments have demonstrated that the lasRNAs of excludons are directly involved in the negative regulation of sense transcripts because elimination of the lasRNA resulted in increased expression of the sense transcript for both the *mogR/lmo0675–lmo0676–lmo0677* operon[10] and the *lmo0605/lmo0606–lmo0607–lmo0608* operon[11]. However, the exact mechanism of excludon asRNA-mediated inhibition of the overlapped ORFs has not yet been determined. One possible mechanism involves selective degradation of the double-stranded RNA that results from hybridization of the overlapping sense and antisense transcripts[41]. In agreement with this, the study of RNase III-mediated digestion in *S. aureus* and other bacteria (see above)[41] identified short RNAs that specifically map to the Anti0677 (*mogR/lmo0675–lmo0676–lmo0677*) excludon locus in *L. monocytogenes*, suggesting that inhibition of the flagellum export apparatus genes occurs by specific, RNase III-dependent cleavage of the double-stranded RNAs that are predicted to form when Anti0677 is expressed. Moreover, northern blot analyses of excludon loci[11] showed that lasRNA transcripts appear as multiple fragments of different sizes, which might be the result of digestion by RNase enzymes. It remains to be determined whether this mechanism or other mechanisms, such as transcription interference or translation inhibition, control excludon regulation.

Concluding remarks

High-throughput sequencing methods have revealed an abundance of asRNAs in microbial genomes, but expression of only a few has been confirmed by independent methods, and even fewer asRNAs have been functionally characterized. Consequently, the functional relevance of asRNAs is debated in the literature, and there have been claims that some asRNAs might be experimental artefacts[8] or might represent transcriptional noise resulting from chance transcription events[33]. Nevertheless, it has also been proposed that because such chance transcription events are metabolically costly, most asRNAs are likely to carry out some function. Accordingly, some asRNAs have been demonstrated to control gene expression, highlighting their pivotal role in the cell. All asRNAs share a common feature — perfect complementarity with their target — but the extreme mechanistic diversity highlighted in the examples mentioned above demonstrates that these transcripts are unlikely to form a uniform group of regulators. With only a handful of examples studied so far, it is too early to establish a dogma for asRNA-mediated regulation. However, the increasing number of functional studies of asRNAs is paving the way for the identification of recurring patterns and the definition of groups of regulatory asRNA according to their mechanism of action.

The excludon paradigm highlights the fact that a single regulatory signal can result in an expression switch between one operon and the neighbouring operon through lasRNAs. In other words, an excludon functions as a genomic toggle, ensuring finely tuned regulation of genes that are encoded at the same genomic locus. So far, two excludons (the *mogR/lmo0675–lmo0676–lmo0677* locus and the *lmo0605/lmo0606–lmo0607–lmo0608* locus) have been characterized in detail. However, the presence of several excludon-like loci in the *L. monocytogenes* genome[11] and the observation that excludons are conserved in *L. innocua*[11], in *Listeria grayi* and *Listeria welshimeri* (as demonstrated by northern blots of the *mogR/lmo0675–lmo0676–lmo0677* locus; N.S. and A. Toledo-Arana, unpublished observations), and in the more distantly related bacterium *Clostridium difficile*[11] indicate that excludon-mediated regulation could be a general mechanism that possibly extends to other organisms outside the genus *Listeria*. Furthermore, the structure of the excludon locus highlights the importance of both genomic context and the spatial distribution of genes along the chromosome for the control of gene expression.

Nina Sesto, Cristel Archambaud and Pascale Cossart are at the Unité des Interactions Bactéries-Cellules, Institut National de la Santé et de la Recherche Médicale (INSERM) U604, Institut Pasteur, and the Institut Scientifique de Recherche Agronomique (INRA) USC2020, Institut Pasteur, Paris F-75015, France.

Omri Wurtzel and Rotem Sorek are at the Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel.

Correspondence to P.C., R.S.
e-mails: pcossart@pasteur.fr; rotem.sorek@weizmann.ac.il

doi:10.1038/nrmicro2934
Published online 24 December 2012

1. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
2. Waters, L. S. & Storz, G. Regulatory RNAs in bacteria. *Cell* **136**, 615–628 (2009).
3. Sorek, R., Kunin, V. & Hugenholtz, P. CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Rev. Microbiol.* **6**, 181–186 (2008).
4. Arraiano, C. M. *et al.* The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.* **34**, 883–923 (2010).
5. Breaker, R. R. Prospects for riboswitch discovery and analysis. *Mol. Cell* **43**, 867–879 (2011).
6. Kortmann, J. & Narberhaus, F. Bacterial RNA thermometers: molecular zippers and switches. *Nature Rev. Microbiol.* **10**, 255–265 (2012).
7. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* **43**, 880–891 (2011).
8. Georg, J. & Hess, W. R. *cis*-antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 286–300 (2011).
9. Thomason, M. K. & Storz, G. Bacterial antisense RNAs: how many are there, and what are they doing? *Annu. Rev. Genet.* **44**, 167–188 (2010).
10. Toledo-Arana, A. *et al.* The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950–956 (2009).
11. Wurtzel, O. *et al.* Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol. Syst. Biol.* **8**, 583 (2012).
12. Stougaard, P., Molin, S. & Nordström, K. RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc. Natl Acad. Sci. USA* **78**, 6008–6012 (1981).
13. Tomizawa, J., Itoh, T., Selzer, G. & Som, T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc. Natl Acad. Sci. USA* **78**, 1421–1425 (1981).
14. Simons, R. W. & Kleckner, N. Translational control of IS10 transposition. *Cell* **34**, 683–691 (1983).
15. Brantl, S. Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr. Opin. Microbiol.* **10**, 102–109 (2007).
16. Guell, M. *et al.* Transcriptome complexity in a genome-reduced bacterium. *Science* **326**, 1268–1271 (2009).
17. Wurtzel, O. *et al.* A single-base resolution map of an archaeal transcriptome. *Genome Res.* **20**, 133–141 (2010).
18. Sharma, C. M. *et al.* The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**, 250–255 (2010).
19. Mitschke, J. *et al.* An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc. Natl Acad. Sci. USA* **108**, 2124–2129 (2011).
20. Mandin, P., Repoila, F., Vergassola, M., Geissmann, T. & Cossart, P. Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res.* **35**, 962–974 (2007).
21. Mraheil, M. A. *et al.* The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. *Nucleic Acids Res.* **39**, 4235–4248 (2011).
22. Nicolas, P. *et al.* Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**, 1103–1106 (2012).
23. Kröger, C. *et al.* The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc. Natl Acad. Sci. USA* **109**, 1277–1286 (2012).
24. Wilms, I., Overlöper, A., Nowrousian, M., Sharma, C. M. & Narberhaus, F. Deep sequencing uncovers numerous small RNAs on all four replicons of the plant pathogen *Agrobacterium tumefaciens*. *RNA Biol.* **9**, 446–457 (2012).
25. Wurtzel, O. *et al.* The single-nucleotide resolution transcriptome of *Pseudomonas aeruginosa* grown in body temperature. *PLoS Pathog.* **8**, e1002945 (2012).
26. Kawano, M., Aravind, L. & Storz, G. An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol. Microbiol.* **64**, 738–754 (2007).
27. Stazic, D., Lindell, D. & Steglich, C. Antisense RNA protects mRNA from RNase E degradation by RNA–RNA duplex formation during phage infection. *Nucleic Acids Res.* **39**, 4890–4899 (2011).
28. Hernández, J. A. *et al.* Identification of a *furA cis* antisense RNA in the *Cyanobacterium Anabaena* sp. PCC 7120. *J. Mol. Biol.* **355**, 325–334 (2006).
29. Rasmussen, S., Nielsen, H. B. & Jarmer, H. The transcriptionally active regions in the genome of *Bacillus subtilis*. *Mol. Microbiol.* **73**, 1043–1057 (2009).
30. Frohlich, K. S., Papenfort, K., Berger, A. A. & Vogel, J. A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. *Nucleic Acids Res.* **40**, 3623–3640 (2012).
31. Vogel, J. *et al.* RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* **31**, 6435–6443 (2003).
32. Sorek, R. & Cossart, P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nature Rev. Genet.* **11**, 9–16 (2010).
33. Raghavan, R., Sloan, D. B. & Ochman, H. Antisense transcription is pervasive but rarely conserved in enteric bacteria. *mBio* **3**, e00156-12 (2012).
34. Callen, B. P., Shearwin, K. E. & Egan, J. B. Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol. Cell* **14**, 647–656 (2004).
35. Andre, G. *et al.* S-box and T-box riboswitches and antisense RNA control a sulfur metabolic operon of *Clostridium acetobutylicum*. *Nucleic Acids Res.* **36**, 5955–5969 (2008).
36. Adhya, S. & Gottesman, M. Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell* **29**, 939–944 (1982).
37. Stork, M., Di Lorenzo, M., Welch, T. J. & Crosa, J. H. Transcription termination within the iron transport-biosynthesis operon of *Vibrio anguillarum* requires an antisense RNA. *J. Bacteriol.* **189**, 3479–3488 (2007).
38. Giangrossi, M. *et al.* A novel antisense RNA regulates at transcriptional level the virulence gene *icsA* of *Shigella flexneri*. *Nucleic Acids Res.* **38**, 362–3375 (2010).
39. Dühring, U., Axmann, I. M., Hess, W. R. & Wilde, A. An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc. Natl Acad. Sci. USA* **103**, 7054–7058 (2006).
40. Lee, E.-J. & Groisman, E. A. An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol. Microbiol.* **76**, 1020–1033 (2010).
41. Lasa, I. *et al.* Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc. Natl Acad. Sci. USA* **108**, 20172–20177 (2011).
42. Lioliou, E. *et al.* Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet.* **8**, e1002782 (2012).
43. Opdyke, J. A., Fozo, E. M., Hemm, M. R. & Storz, G. RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J. Mol. Biol.* **406**, 29–43 (2011).
44. Opdyke, J. A., Kang, J.-G. & Storz, G. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Mol. Biol.* **186**, 6698–6705 (2004).
45. Tramonti, A., De Canio, M. & De Biase, D. GadX/GadW-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the *gadY–gadW* divergent promoters and identification of four novel 42bp GadX/GadW-specific binding sites. *Mol. Microbiol.* **70**, 965–982 (2008).
46. Sakurai, I. *et al.* Positive regulation of *psbA* gene expression by *cis*-encoded antisense RNAs in *Synechocystis* sp. PCC 6803. *Plant Physiol.* **160**, 1000–1010 (2012).
47. Sayed, N., Jousselin, A. & Felden, B. A *cis*-antisense RNA acts in *trans* in *Staphylococcus aureus* to control translation of a human cytolytic peptide. *Nature Struct. Mol. Biol.* **19**, 105–112 (2011).
48. Arnvig, K. B. & Young, D. B. Identification of small RNAs in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **73**, 397–408 (2009).
49. Mandin, P. & Gottesman, S. Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J.* **29**, 3094–3107 (2010).
50. Papenfort, K. *et al.* Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol. Microbiol.* **74**, 139–158 (2009).
51. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature Rev. Micro.* **6**, 613–624 (2008).

Acknowledgements
The authors thank J. Mellin for insightful discussions. This work was supported by the Pasteur–Weizmann programme (grants to P.C. and R.S.), the European Research Council (Advanced Grant 233348 to P.C. and Starting Grant 260432 to R.S.), the French Agence Nationale de la Recherche (grants Bacregrna 09-BLAN-0024-02 to P.C. and BacNet 10-BINF-02-01 to P.C. and N.S.), the Fondation Le Roch Les Mousquetaires (P.C.), the Fondation Louis-Jeantet (P.C.), the German Minerva Foundation (R.S.), the Deutsche Forschungsgemeinschaft (a German–Isaraeli Project Cooperation grant to R.S.), the French Institute National de la Recherche Agronomique (a fellowship to N.S.) and the Azrieli Foundation (a fellowship to O.W.).

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Pascale Cossart's homepage:
http://www.pasteur.fr/ip/easysite/pasteur/fr/recherche/departements-scientifiques/biologie-cellulaire-et-infection/unites-et-groupes/unite-des-interactions-bacteries-cellules
Rotem Sorek's homepage:
http://www.weizmann.ac.il/molgen/Sorek/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF