UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in Scienze Biologiche e Molecolari

XXV Ciclo

Post-transcriptional regulation in Gram-negative bacteria

Francesco Delvillani

PhD Thesis

Scientific tutor: Dr. Federica Briani

Academic year: 2012

SSD: BIO/18; BIO/19

Thesis performed at Bioscience Department

Contents

ABSTRACT

Post-transcriptional regulation plays a pivotal role in gene expression control as it ensures a fast response to environmental changes. This is particularly important in unicellular organisms that are exposed to environmental changes. During my PhD, I studied different aspects of this type of regulation in two Gram-negative bacteria: *Escherichia coli* and the opportunistic pathogen *Pseudomonas aeruginosa*. In *E. coli* I investigated the role of the ribosomal protein S1 in the interplay between translation and mRNA decay. In particular we showed that S1 not associated to 30S can sequester and protect mRNA from RNase E, the main endonuclease in *E. coli*. Another aspect that I considered in my research has been regulation by RNA molecules in *P. aeruginosa*. In the last decade, it has become clear how RNA-based regulatory mechanisms are important in controlling bacterial gene expression. However, little is known about it in this relevant human pathogen. By RNA deep sequencing we identified more than 150 novel candidate sRNAs in the *P. aeruginosa* strains PAO1 and PA14, which differ in virulence degree. We confirmed by Northern blotting the expression of 52 new sRNAs, substantially increasing the number of known sRNAs expressed by this bacterium. In this context we developed a genetic screen for the identification of genes post-transcriptionally regulated by RNA determinants and applied this system to the search of RNA thermometers (RNATs), *i.e.* mRNA determinants that couple translation with temperature changes. We identified four putative RNATs and validated two of them in *E. coli*. Interestingly, the two are located upstream of genes previously implicated in *P. aeruginosa* pathogenesis, namely *dsbA* and *ptxS*. *ptxS* RNAT was validated also in *P. aeruginosa* and represents the first RNAT ever described in this bacterium.

PART I

Bacteria are highly adaptable organisms able to survive in a wide range of environments with variable temperature, pH, availability of water, salts and nutrients. This is due to high metabolic versatility and efficient gene regulation that allow them to respond effectively to environmental changes by accordingly modulating cell physiology.

Genes can be regulated at many stages, from transcription initiation to protein activity and degradation. The early elucidation of regulation mechanisms of *Escherichia coli* lactose operon and lambda phage lysogenic state, both based on repressors modulating transcription initiation, prompted the idea that gene expression would be essentially regulated by such mechanism. Conversely, successive research has enlarged the repertoire of bacterial regulatory strategies, showing that, albeit transcription initiation is indeed regulated at many promoters, further steps of gene expression are also targeted by regulators. Moreover, it has been found that the chemical nature of regulators is not limited to proteins, as RNAs and small molecules (i.e. modified nucleotides) control relevant aspects of bacterial physiology.

My research as a PhD student has addressed different aspects of translation initiation regulation and interplay between translation and mRNA decay in two Gram negative bacteria, *Escherichia coli* and *Pseudomonas aeruginosa*.

Escherichia coli is a Gram-negative facultative anaerobic bacterium that belongs to the class of Gammaproteobacteria. It is commonly found in the lower intestine of humans and other warm-blooded organisms. The first complete genome sequence of *E. coli* was published in 1997; the sequenced strain (*E. coli* K-12 MG1655) owns a $4.6x10^6$ bp genome with 4288 annotated open reading frames (ORFs) [1]. Sequencing of other *E. coli* strains showed that this species has a genome ranging from $4.6x10^6$ bp to $5.5x10^6$ bp. *E. coli* is the most widely studied prokaryote and the best known model organism. It is also widely used in the field of biotechnology and as host for heterologous gene expression. However, though profusely studied, about 30% of *E. coli* genes have unknown function [2].

Most *E. coli* strains are harmless, but some can cause from mild to severe infections. Eight pathovars have been described and their mechanism of pathogenesis studied. Six of them are diarrheagenic and the other two cause extraintestinal infections. The EPEC strain is the major cause of infection in developing countries. This strain acquired a particular pathogenicity island (PAI), named LEE (locus of enterocyte effacement). This element encodes a type III secretion system (T_3SS) used to translocate bacterial effectors into the host cells. The genome of pathogenic strains may be $1x10^6$ bp larger than that of commensal strains. They usually contain PAIs and plasmids

coding for genes involved in the attachment, in the virulence and in the proliferation into the host cells [3].

Pseudomonas aeruginosa is a highly adaptable Gram-negative bacterium, which thrives in a broad range of ecological niches. It is usually found in the soil and is able to grow also in hypoxic environment. Moreover, it can infect organisms as different as nematodes, mammalians and plants. This bacterium owns a relatively big (6-7 Mb), GC-rich genome with around 6000 predicted open reading frames. In human, *P. aeruginosa* behaves as an opportunistic pathogen infecting wounds, burns and medical devices. In cystic fibrosis patients *P. aeruginosa* is the most common pathogen that infects the respiratory tract and it is the main responsible of respiratory failure and mortality among these patients. This bacterium produces a large number of virulence factors and is intrinsically resistant to several antibiotics. The resistance of *P. aeruginosa* to antibiotics depends on different factors. First of all, *P. aeruginosa* has a low outer membrane permeability compared to other bacteria such as *E. coli*. Moreover, this bacterium has an efficient system of efflux pumps that ensures the export of antibiotics out of the cytoplasm. Finally, it has been reported that 37% of CF patients chronically infected by *P. aeruginosa* carry hypermutator strains. These strains harbor mutations in DNA repair genes and are thus subject to mutation accumulation. (Multi)Resistant strains are selected by the intensive usage of antibiotics in these patients [4].

mRNA decay in Bacteria

The amount of an mRNA depends on the balance between the rate of its transcription and degradation. In *E. coli,* mRNA half-lives range from less than 30 seconds to more than 20 minutes. This variability depends on intrinsic features of RNA molecules, such as the presence of specific sequences and secondary structures. For instance, the presence of Rho-independent terminators at the 3'-end may prevent decay by exonucleases that are inhibited by double stranded (ds) RNA [5, 6]. On the other hand, polyadenylation at the 3'-end has a destabilizing role, giving the graft to exonucleases [7]. Translation rate may also modulate transcript half-life: it is commonly accepted that mRNA efficiently translated are more stable than untranslated transcripts, suggesting that ribosomes may play an mRNA protective role [8]. Finally, specific regulators, either RNAs or proteins, may affect mRNA stability in response to different stimuli.

Most of the studies on mRNA decay in bacteria have been performed on two model organisms: the Gram-negative *E. coli* and the Gram-positive *Bacillus subtilis*. These studies have shown that relevant differences exist between the two bacteria in the RNA degradation strategy.

In *E. coli*, the main pathway of mRNA decay begins with a cleavage by RNase E that usually occurs in the 5'-untranslated region (5'-UTR) and removes the translation initiation region (TIR) of the mRNA. This step corresponds to mRNA functional inactivation, as the transcript is no longer translatable, and it is usually the rate limiting step in decay pathways, because RNAs with a triphosphate at the 5'-end are not good substrates for RNase E (see below). Further mRNA fragmentation due to other RNase E-mediated endonucleolytic cuts usually ensues. The RNA fragments are degraded by 3'-to-5' exonucleases such as PNPase and RNase II. Finally, oligoRNase digests short oligoribonucleotides, thus fleshing the cellular nucleotide pool out.

Albeit this is commonly described as the most common decay pathway in *E. coli*, it is worth mentioning that this notion is based on the study of the decay of few model mRNAs and that, despite of this, mRNAs that are not degraded through the strategy outlined above have been described. This is the case for instance of the monocistronic *rpsO* mRNA. Its degradation starts with an endonucleolytic cleavage by RNase E in the 3'-UTR (at the M2 site). This cut removes the 3'-terminal RNA hairpin from the mRNA, thus converting it into an ideal substrate for the exonuclease PNPase [9]. M2 is not the only RNase E site in the *rpsO* mRNA. In fact *rpsO* transcript with an M2 mutation is still degraded through an RNase E-dependent decay. Moreover, in this mutant, a 3'-end polyadenylation-dependent degradation pathway seems to be enhanced. In fact, the absence of PAPI polyadenylpolymerase I (PAPI), which does not affect *wt rpsO* mRNA half-life, stabilizes the M2 mutant transcript [10]. *rpsO* mRNA decay is an example showing how an RNA can enjoy multiple and interconnected degradative pathways; regulation of the choice among different decay pathways remains still unclear.

B. subtilis adopts different RNA degradation strategies. It lacks RNase E, the main endonuclease of Gram-negatives, and owns two 5'-to-3'exonucleases, RNase J1 and J2, which are absent in *E. coli* and other Gram-negatives and present in other Gram-positives. RNase J1 and J2 are also endowed with endonucleolytic activity. Decay in this bacterium begins with a cleavage by RNase J1/J2 within the body of the mRNA. The two fragments generated by the cut are the substrate for the 3'-to-5' and 5'-to-3' exonucleolytic action of PNPase and RNase J1/J2, respectively [11]. Exonucleolytic degradation from the primary transcript 5'-end by RNase J1/J2 represents an alternative degradation pathway [12].

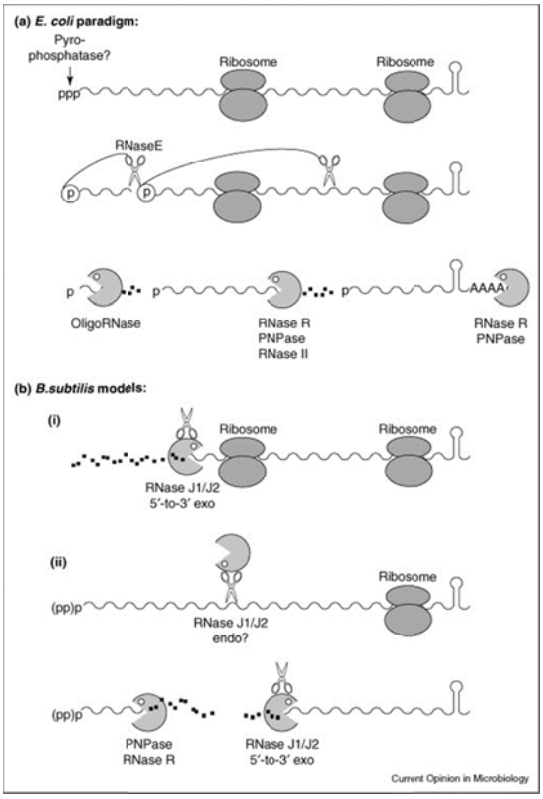


Figure 1. Principal mRNA decay pathways in *E. coli* and *B. subtilis*. **(a)** *E. coli* paradigm. Scissors symbol indicates internal cut of mRNA by RNase E. "Pacman" symbol refers to the major 3′ exonucleases (RNase II, PNPase, RNase R or OligoRNase). RNA fragments with secondary structures, such as Rho-independent terminator hairpins, need polyadenylation before being degraded. OligoRNase degrades oligoribonucleotides of lenght ranging from 2 to 5 nt **(b)** *B. subtilis* models. **(i)** RNA fragments are degraded from their 5'-end by the exoribonuclease RNase J1/J2. The double exonucleolytic and endonucleolytic activities of these enzymes are indicated by the combined Pacman/scissors symbols. **(ii)** The first endonucleolytic cut, by RNase J1/J2, occurs within the body of the mRNA. The two sides of the cleavage provide a 3'-OH and a 5'-Pi end for the attack of PNPase and RNase R or RNase J1/J2 respectively. From [12].

E. coli enzymes involved in mRNA decay

Endoribonucleases

These enzymes cut RNAs at internal positions and play a crucial role in both mRNA decay and processing and in the maturation of ribosomal and transfer RNAs.

RNase E is the main endonuclease in *E. coli* mRNA decay. The monomer (118 kDa) is encoded by the essential gene *rne*. It was initially identified by a temperature-sensitive mutation (*rne-3071*) as the enzyme able to process the 9S RNA into the 5S rRNA [13]. Later, a thermosensitive

mutant (*ams-1*)[14] that exhibited altered bulk mRNA stability was discovered to have a defect in RNase E, thus outlining the relevant role of this enzyme in the mRNA decay.

The protein is catalytically active as a homotetramer (dimer of dimers). The monomer can be structurally divided into two distinct halves: the globular amino-terminal (NTH; residues 1-529) and the carboxyl-terminal (CTH; residues 530–1061) domains. The NTH harbors the catalytic domain and can be further divided into a large and a small domains linked by a short Zn-link domain that coordinates a Zn^{2+} ion that stabilizes the RNase E dimer [15]. The small domain is involved in the stabilization of the tetramer [16]. The CTH region is needed to interact with other proteins [17, 18] and to form a high molecular weight complex named RNA degradosome. The RNA degradosome is constituted by RNase E, the exoribonuclease PNPase, the helicase RhlB and the glycolytic enzyme enolase [19-21]. The complex is able to degrade hairpin structures that usually protect the 3'-end from exonucleolytic decay [22]. RNase E cleaves the phosphodiester backbone of the RNA, generating 3'-OH and 5'-monophosphate termini [23]. It shows a preference for single stranded A/U-rich RNA sequences [24, 25]. Moreover, because of the characteristics of the central channel, the catalytic site is poorly accessible for RNAs with a triphosphate 5'-end, thus rationalizing the observation that RNAs with monophosphorylated 5'-termini are preferred substrates [26].

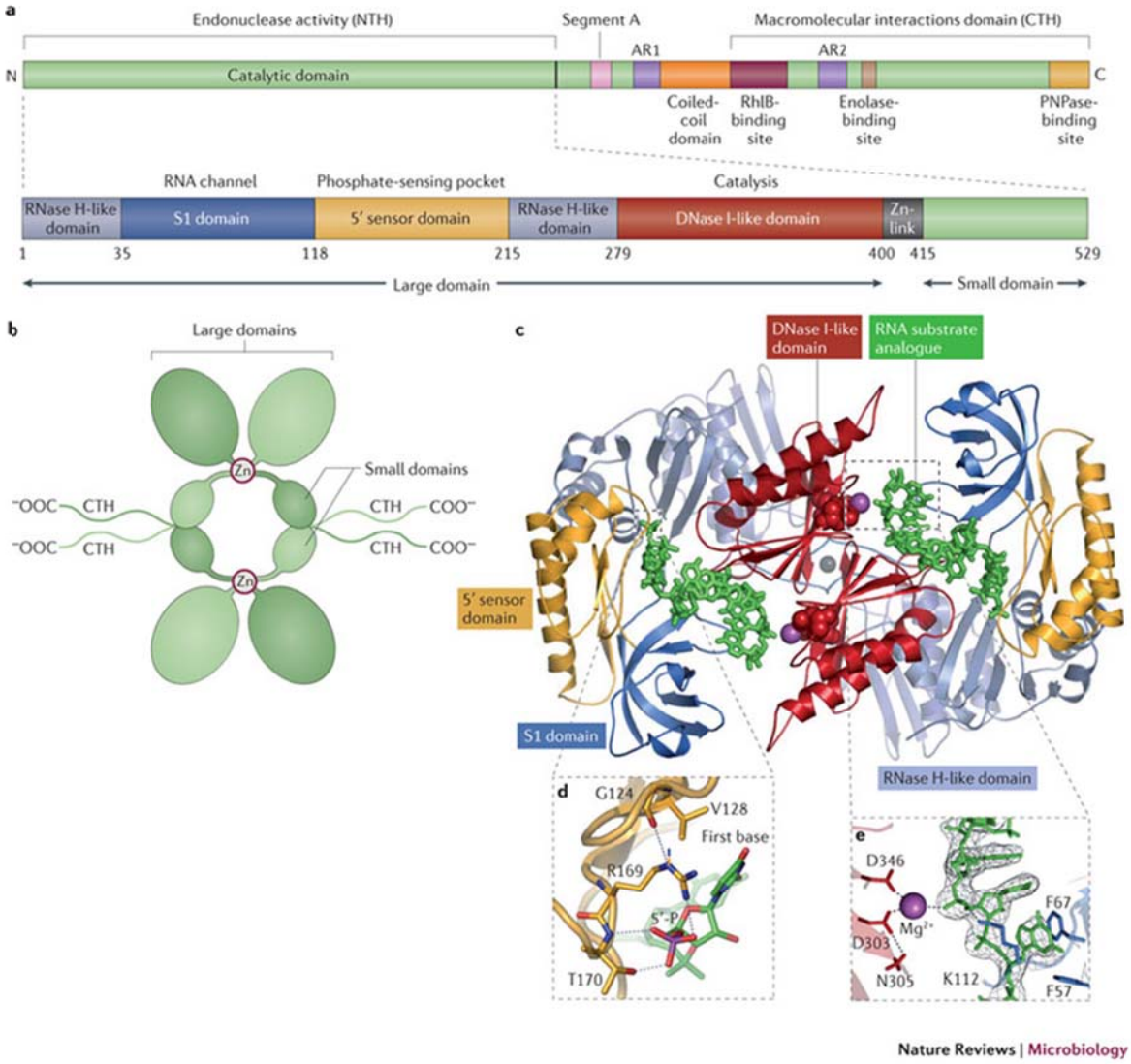


Figure 2. (a) Schematic representation of the primary structure of an RNase E monomer. The NTH is the catalytic domain and can be divided into two globular domains that are further divided into several functional subdomains. The carboxy-terminal half (CTH) is the domain involved in the interaction with other molecules and contains binding sites for various interacting proteins. The two boxes AR1 and AR2 are the arginine-rich segments that might have a role in the RNA binding. **(b)** *In vivo*, RNase E works as a tetramer formed by two dimers. Each RNase E monomer is colored in light or dark green. Monomer interaction is stabilized by both the dimer interface and coordination of a Zn^{2+} ion through the Zn-link. Interactions among the four small domains stabilize the tetramer. **(c)** Crystal structure of an RNase E dimer together with an analogue substrate. The second dimer lies below the plane of the drawing. The grey dot in the center is a Zn atom in the Zn-link. **(d)** Enlargement of the 5′ sensor subdomain of the NTH domain, showing residues Arg169 and Thr170, that interact with the terminal phosphate of the substrate. Val128 stacks on the first base of the substrate, and Gly124 stabilizes the side chain of Arg169. **(e)** Enlargement of the active-site metal ion. Residues Asp303 and Asp346 chelate the metal ion and residues Phe57, Phe67 and Lys112 contact the substrate. From[27].

Some evidence suggests that RNase E is associated with the inner membrane [28, 29]. It has been remarked that if RNase E is immobilized at the periphery of the cell, translation and decay may actually occur in different cell compartments at least for some RNAs, thus questioning current models asserting that an mRNA half-life depends on competition between ribosomes and mRNA decay machinery [27, 30, 31].

RNase E regulates its own mRNA abundance by cleaving it in the 5'-UTR [32]. In addition, RNase E activity is modulated by the interaction with different proteins. The ribosomal protein L4 inhibits its activity, suggesting a tight cooperation between translation and decay machinery [33]. Moreover, recent findings suggest that RNase E interactors may exist that recruit the protein to specialized decay pathways. This can be the case of RapZ (RNase adaptor protein for sRNA GlmZ), a protein involved in the decay of the small RNA *GlmZ*. RapZ is an RNA-binding protein that interacts with and recruits RNase E to *GlmZ* [34].

RNase G, encoded by *rng*, is homologous to RNase E NTH and is involved (together with RNase E) in the maturation of 16S rRNA. In an *rng* null mutant, the 16.3S rRNA, a precursor of 16S, accumulates [35]. This enzyme is also involved in mRNA turnover [36] and, in its absence, the half-life of some mRNAs increases [37]. Overexpression of this protein can complement the absence of RNase E, thus suggesting that the two proteins are endowed with partially overlapping functions [38].

RNase Z is a conserved zinc-dependent endonuclease. Originally identified for its involvement in 3'-end maturation of tRNAs [39], RNase Z seems to play a role also in the decay of some *E. coli* mRNAs [40].

RNase III, encoded by the non-essential *rnc* gene, is a dimer of two 25 kDa subunits; mutants lacking the catalytic activity do not grow at ≥ 45 C° [41]. It has affinity for dsRNA on which makes staggered cuts [42, 43]. This nuclease contributes to the maturation of 16S and 23S rRNAs and modulates the stability of several bacterial and phage transcripts [44]. In many cases its cleavage triggers mRNA destabilization, as occurs for instance for its own *rnc* mRNA or in the *pnp* autoregulation circuit [45, 46]. For other mRNAs, the cleavage does not alter the half-life, but rather modulates the translation efficiency[44].

RNase III plays an important role in the degradation and/or processing of sRNA-mRNA hybrids. For instance, it has been observed that the sRNA MicA is cleaved by this enzyme when coupled with its mRNA target *ompA* [47]; moreover, this enzyme is involved in the processing of RNAI (a small non coding RNA that regulates ColE1 plasmid copy number [48]).

In Eukaryotes, belong to RNase III family Drosha and Dicer, two enzymes important for the maturation of microRNAs, small non coding RNAs with regulatory properties. Drosha is the enzyme that produces pre-microRNA (roughly 70nt long) from longer transcripts. Dicer processes the pre-microRNAs to the typical mature microRNAs (20-25nt) able to bind their targets genes [49].

RNase P is a ribozyme composed by a small protein of about 15 kDa (encoded by *rpnA* gene) and a non-coding (nc) RNA roughly 400 nt long (encoded by *rpnB* gene) with a stoichiometry of 1:1 [50]. The RNA component shows catalytic activity, but it requires the protein subunit to work efficiently [51]. This enzyme is primarily involved in the maturation of the 5'-end of pre-tRNA to generate tRNA [52]. However, low efficiency cleavages by RNase P have been observed in several polycistronic mRNAs [53, 54].

Exoribonucleases

All eight known *E. coli* exonucleases degrade the RNA in 3'-to-5' direction; three of them (PNPase, RNase II and RNase R) play crucial roles in mRNA decay. RNase II and R (encoded by the *rnb* and *rnr* genes, respectively) degrade the RNA by an hydrolytic reaction that releases nucleoside monophosphates (NMPs), whereas PNPase (encoded by *pnp*) is a phosphorolytic nuclease and releases nucleoside diphosphates (NDPs). None of these three exonucleases is essential, suggesting a certain degree of redundancy. However, *pnp* null mutations are synthetically lethal with mutations inactivating either RNase II or R. Conversely, double *rnb rnr* mutants are viable. This implies that the two hydrolytic RNases are not fully interchangeable and PNPase may obviate their lack. For a review see [55].

The enzymes of the RNase II-family of exoribonucleases are present in all kingdom of life [56, 57]. *E. coli* **RNase II** is the prototype of the RNase II-family. This enzyme is composed by four domains, namely two cold shock domains (CDS1-CDS2) at the N-terminal followed by an RNB domain and an S1 domain at the C-terminal. The RNB is the catalytic domain, whereas the others are responsible for recognizing and binding RNA [58]. The interaction between RNA and RNase II involves two different portions of the enzyme: the anchoring region (constituted by the three binding domains CSD1-CSD2-S1) and the catalytic region. The shortest RNA fragment that retains both interactions and that is processively degraded by RNase II is 10 nt long. The catalytic region creates a hole that allows the access only to single-stranded RNA. This provides a structural basis for the observation that RNase II is inhibited by secondary structures. Detailed description of the structure has been reviewed in [59]. RNase II is responsible for the 90% of the exoribonucleolytic activity in crude extract of exponentially growing *E. coli* cells. This suggests a predominantly

hydrolytic activity in this organism. In contrast, in *B. subtilis*, that lacks an RNase II homologue, the decay is mostly phosphorolytic [60].

RNase II degrades the poly(A) tail at the 3'-end of messenger RNAs [61]. In some cases, this activity leads to a stabilization of the transcripts. For instance, the poly(A) tail of *rpsO* mRNA is needed for PNPase degradation and is removed by RNase II. Marujo et al. have demonstrated that *rpsO* mRNA is destabilized in a strain deficient for RNase II, whereas is stabilized by its overexpression. This holds true only when poly(A) polymerase is active [62].

RNase R is also a processive exoribonuclease belonging to the RNase II family. This enzyme, differently from RNase II and PNPase, can degrade structured RNA if a 3' single-stranded overhang is present [6]. At low temperature, *rnr* defective mutants show a defect of growth. This could be linked to the absence of the hydrolytic activity of this RNase on structured RNA that becomes particularly stable at low temperature. This enzyme is involved in the maturation of stable RNA and degrades aberrant tRNA and rRNA fragments, whose accumulation could be potentially harmful [6, 63, 64]. The protein is induced by cold shock and during stationary phase. It is been implicated in the maturation of SsrA/tmRNA and in the decay of *ompA* mRNA [65, 66].

PNPase catalyses the *in vivo* phosphorolytic degradation of RNA and the reverse reaction of heteropolymeric tail synthesis at the 3'-end of RNA [67]. The degradative reaction occurs through a processive mechanism in which the enzyme does not dissociate from the substrate until it has reached a length of less than 20 nucleotides [68, 69]. The enzyme is highly conserved in bacteria and is also found in the mitochondria and chloroplasts of some higher eukaryotes, while it is missing in Archea and Fungi. In human cells PNPase has been found in the intermembrane space of mitochondria [70]. Recently it has been observed that a missense mutation in PNPT1 gene (encoding the human PNPase) reduces the 5S rRNA import into this organelle [71]. The presence of PNPase in the cytoplasm and its involvement in the mRNA and miRNA decay have been also claimed [72].

PNPase is a homo-trimer, whose subunits generate a central channel [73, 74]. Each monomer is 711 aa long and is constituted at the N-terminus by two RNase PH domains linked by an all alpha-helix domain. The RNA binding domains KH and S1, both necessary for binding the RNA, are located at the C-terminus. The deletion of either KH or S1 or of both domains reduce the affinity for RNA and, in turn, the enzymatic activity [75-77]. In *E. coli*, PNPase has been found free or associated with the RNA helicase RhlB; it is also a component of the RNA degradosome [20, 21, 78].

The *pnp* gene is not essential at 37° but becomes essential at low temperature [79, 80]. Interestingly, it has been recently found that PNPase activity can be modulated (at least *in vitro*) by the interaction with small molecules such as ATP, ppGpp, c-diGMP and citrate [81-83 84] suggesting

that the metabolic state of the cell can impact on the activity of this enzyme. Like the RNase II, PNPase seems to be able to protect mRNA from degradation in specific conditions. Several mRNAs show a PNPase-dependent stabilization upon over-expression of the ribosomal protein S1 [85] and some sRNAs show decreased stability in a *pnp* defective mutant [86].

It has been demonstrated that PNPase is able to lengthen or shorten the 3'-end of single stranded DNA; interestingly, this activity is conserved between two species of bacteria evolutionarily distant as *E. coli* and *B. subtilis*. Although the physiological role of this activity is still unclear, this and other evidence link PNPase to DNA repair machinery (T. Carzaniga and G. Dehò, unpublished) [87, 88].

PNPase regulates its own expression by modulating *pnp* transcript stability through a complex mechanism that involves RNase III and E (Fig. 3) [89, 90].

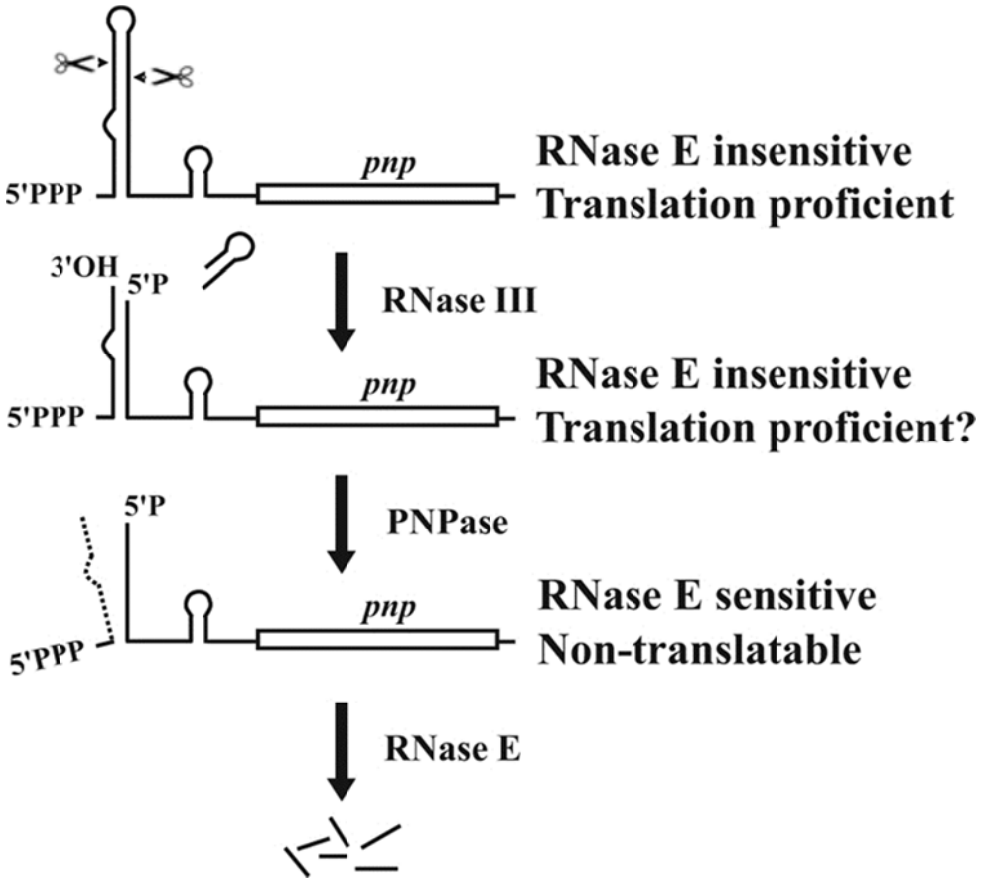


Figure 3. Scheme of PNPase autoregulation. RNase III (scissors) makes a staggered cleavage in the 5'-UTR of *pnp* mRNA leaving two still paired RNA fragments. PNPase degrades the RNA fragment upstream of the cleavage site. After this step, the *pnp* mRNA becomes sensitive to RNase E attack and quickly degraded. From [90].

OligoRNase. All known *E. coli* exonucleases leave RNA fragments longer than 2 nucleotides, whose accumulation is detrimental to the cell [91]. OligoRNase, which degrades these short fragments (2-5 nucleotides) to mononucleotides, is the only *E. coli* essential exoribonuclease.

Other activities involved in mRNA degradation

PAPI (polyadenyl polymerase I) polymerizes ATP into poly(A) tails at the 3'-end of messengers [92]. Many RNAs carry, at their 3'-end, the stable stem loop of Rho-independent terminators. This structure hampers the attack by 3'-to-5'exoribonucleases, whereas the poly(A) tail provides an entry point for RNases. Microarray analysis of RNA extracted from *E. coli* exponential cultures showed that more than 90% of genes is subject to PAPI-dependent polyadenylation [93]. Recently, it has been suggested that PAPI is also necessary to maintain a balanced level of tRNAs. In fact it has been demonstrated that when PAPI level increases, tRNAs become polyadenylated and this leads to a decrease of aminoacylated tRNAs and cessation of protein synthesis. This phenotype is at least partially rescued when the enzymes involved in the maturation of the tRNA 3'-end are lacking [94].

Hfq protein is a homohexameric RNA chaperone [95]. It interacts, stabilizes and promotes the pairing between RNAs. Moreover, it can interact with poly(A) tails and increase the processivity of poly(A) polymerase [96]. The physiological role of this protein in sRNA-based regulation will be discussed below.

RNA pyrophosphohydrolase (RppH) converts the 5′-end of primary transcripts from 5'-triphosphate to 5'-monophosphate, making them more susceptible to the attack of RNase E [97]. In fact, several transcripts are stabilized in the absence of *rppH* [98]. This step has been functionally related to the eukaryotic mRNA decapping.

Translation initiation regulation in Gram-negative bacteria

The bacterial ribosome

The bacterial ribosome is constituted by a small (30S) and a large (50S) subunit that join together to form the monosome (70S). The association of 30S and 50S generates three sites within the ribosome (site A, P and E). Site A is the entry point of tRNA and is occupied by aminoacyl-tRNA (tRNA loaded with the cognate aminoacid), site P is occupied by peptidyl-tRNA (tRNA carrying the growing polypeptidic chain) and site E is the exit site for the uncharged tRNAs.

Three ribosomal RNAs (rRNAs) together with 54 proteins make up the ribosomal particle. The rRNAs are transcribed from ribosomal operons that include also tRNAs genes. In *E. coli* there are seven ribosomal operons (*rrnA, rrnB, rrnC, rrnD, rrnE, rrnG* and *rrnH*), which have very similar organization and are transcribed by two tandem promoters [99]. The long primary transcript undergoes a processing reaction that produces the 5S, 16S and 23S rRNAs. The 5S and 23S form the large subunit, whereas the 16S is a component of the small subunit.

In order to have balanced expression of ribosomal components, the expression of r-proteins and rRNA is co-regulated through multiple mechanisms. Some r-proteins participate in an antitermination complex that assembles on the nascent *rrn* transcript, thus increasing transcription elongation efficiency [100]. In the absence of free r-proteins, premature termination on *rrn* will occur, decreasing the rRNA synthesis. Conversely, when free r-proteins are present, they can participate in the antitermination complex, thus increasing rRNA synthesis. The physiological meaning of this mechanism of regulation is to modulate *rrn* operons transcription according to ribosomal protein availability.

Reciprocal regulation also occurs, as r-proteins translation depends on rRNA availability. For instance, ribosomal protein L4 interacts with its polycistronic mRNA (S10), inhibiting the expression of the whole operon [101]. L4 is a primary ribosomal protein (i.e. it can bind the rRNA in the absence of any other pre-assembled ribosomal protein) and has higher affinity for rRNA than for its own transcript. Thus, L4 will bind its own mRNA and repress its translation only in the absence of free rRNA. Some ribosomal proteins repress translation by entrapping the ribosomal subunit on their mRNA. An example comes from the α operon regulation. This operon encodes the ribosomal proteins S13, S11, S4 and L17 and the α subunit of RNA polymerase core enzyme. The α operon transcript can exist in two conformations in equilibrium between each other, one translationally active and one inactive. Both forms are able to bind the 30S to generate a binary complex (mRNA:30S). The active binary complex can be bound by initiator tRNA and form a normal ternary complex (mRNA:30S:$tRNA^{fMet}$) that will allow translation. The inactive binary complex binds S4 to form an entrapment ternary complex (mRNA:30S:S4) unable to start translation and also to be reconverted in the active form. Thus, S4 not assembled in the 30S modulates the equilibrium between active and inactive forms of its own mRNA (reviewed by [102]).

Translation initiation

Translation initiation is the rate-limiting step during protein synthesis. The first step of translation is the interaction between a sequence at the 3'-end of 16S rRNA and the Shine-Dalgarno region (SD), a polypurinic, 6 nt long mRNA sequence located 6 to 9 nt upstream of the start codon.

The ribosomal protein S1 participates in the early interaction between the 30S subunit and the mRNA by unwinding mRNA secondary structures and allowing efficient ribosome association [103]. After initial interaction, the initiation factors (IFs) are involved in the accommodation of the start codon in the ribosome P-site; in particular, IF1 binds to 30S and prevents the association of 50S and the entering of aminoacyl tRNA. IF2 binds the aminoacylated initiator tRNA (fMet-tRNA^{fMet}) and 30S and, by the hydrolysis of GTP, promotes ribosomal subunits assembly. IF3, which unlike IF1 and 2 is not universal in Bacteria, in *E. coli* is necessary for initiation complex efficient formation. Once the ribosomal subunits and the fmet-tRNA at the P-site have been assembled, the A-site is ready to accept the next aminoacylated tRNA that arrives in a ternary complex with the elongation factor Tu (EF-Tu) and GTP. GTP hydrolysis promotes aa-tRNA accommodation that positions it in the peptidyl transferase center (PTC) on the 50S subunit, where the peptide bond is catalysed. The process ends with hydrolysis of GTP by EF-G that drives the translocation of uncharged tRNA^{fMet} to the E-site, where it can be ejected, and the arrival of the peptidyl-tRNA to the P-site. The A-site is free and ready for the next round of elongation [104].

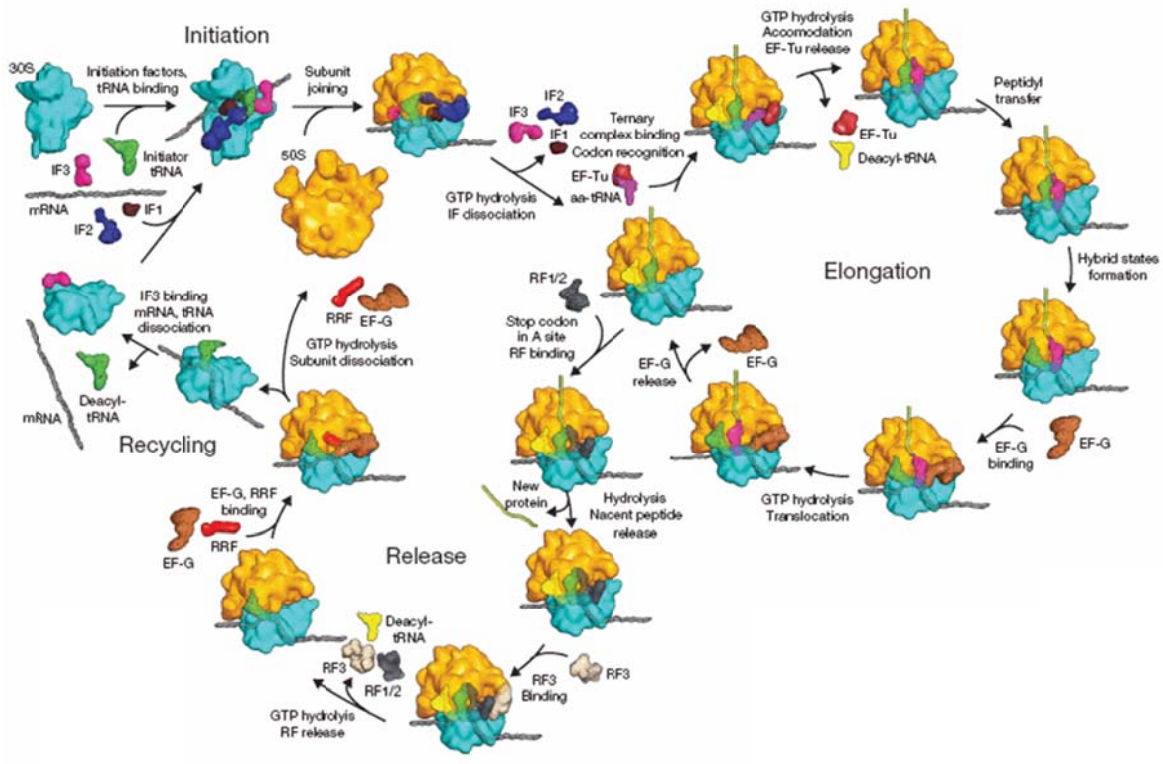


Figure 4. Scheme of the translation steps. 30S in blue; 50S in orange; IF, initiation factors; EF, elongation factors; RF, release factors; aa-tRNA, aminoacyl-tRNA From [104].

Altered translation specificity of modified ribosomes

Modifications in ribosome composition can deeply affect translation initiation, by changing ribosome specificity in mRNA recognition. The content of both the rRNA and the r-protein

components can be modified. For instance, it was shown that in the presence of kasugamycin (Ksg), 61S ribosomal particles, which invariably lack six proteins (S1, S2, S6, S12, S18 and S21) and have a reduced content of other five, are formed. The 61S ribosome is able to translate leaderless mRNAs (starting with a 5'-terminal AUG), whereas translation initiation of canonical mRNA (i.e. with a 5'-UTR) is inhibited [105, 106].

On the other hand, in stress conditions due to translation inhibition or in the presence of high concentration of (p)ppGpp, the expression of the toxin MazF induces the formation of ribosomal particles with a processed 16S RNA. This protein is part of a toxin-antitoxin (TA) system. These systems are usually constituted by two proteins encoded in the same operon. During normal growth, both are expressed and the antitoxin inhibits the toxin. When expression of the TA operon is impaired, as the antitoxin is more unstable than the toxin, the toxin is no longer inhibited and can exert its deleterious effect. Several toxins are RNases with low sequence specificity. Among them, RNA interferases (such as MazF) can attack RNA regardless of its association with the ribosome. Amitai et al. demonstrated that after MazF induction, about 10% of the genes were selectively translated and that MazF is directly responsible for the selective translation of these genes [107, 108]. In fact it was demonstrated that this toxin can cleave 16S rRNA and remove 43 nt at the 3'-end, where the anti-SD lies. Moreover, it was observed that overexpression of MazF induces a cut upstream of the AUG start codon of mRNAs generating transcripts without a SD sequence (leaderless mRNA). *In vitro* translation assays have shown that, unlike canonical mRNAs, these leaderless are translated by ribosomes containing the cleaved 16S ($70S\Delta^{43}$). Interestingly, part of these genes are involved in the cellular response to MazF induced stress. Thus these "*stress-induced*" ribosomes can distinguish between canonical and leaderless mRNA, translating only the latter. This may represent a novel mechanism of gene regulation based on the formation of "programmed ribosomes" that can specifically translate only selected modified mRNAs [108].

These two examples refer to translational stress conditions triggered by either specific inhibitors or severe starvation; it would be interesting to analyse whether specialized ribosomes may exist also in physiological conditions.

Interplay between translation and mRNA decay

Transcription, translation and mRNA decay are interconnected processes. It is commonly accepted that "naked" mRNAs are less stable than mRNAs efficiently translated because ribosomes and RNases may compete for binding the messenger. The functional lifetime of an mRNA is the time span during which an mRNA can be translated, whereas the physical lifetime of an mRNA molecule is the time from its synthesis to its degradation. When a transcript is inactivated by a

nucleolytic event removing the TIR region, the functional and physical lifetimes are identical. On the other hand, functional inactivation may also be a separate event. For example, the *thrS* mRNA of *E. coli*, encoding threonyl tRNA synthetase (ThrS), can be non-nucleolytically inactivated by the binding of ThrS protein itself that competes with the ribosome for binding [109].

The interplay between RNA decay and translation has been the object of my research in the first year my PhD. In particular, I analyzed the role of ribosomal protein S1 in this phenomenon.

S1 ribosomal protein

S1, encoded by the essential gene *rpsA*, is the biggest ribosomal protein (577 aa, 61kDa). Unlike many other r-proteins, that are positively charged and seem to be completely titrated in the cells by the interaction with the ribosome, S1 has an acid isoelectric point and can be found associated to 30S or free in the cytoplasm [110, 111]. S1 consists of six non-identical repetitions of a domain (S1 domain) that has been found in other RNA-binding proteins such as PNPase [16, 112, 113]. The first two N-terminal S1-domains form a globular portion that is responsible for the interaction with S2 and the ribosome [114]. The central and C-terminal regions form an elongated domain involved in mRNA binding. S1 has affinity for ssRNA with a preference for A/U-rich sequences that usually lie upstream of the Shine Dalgarno in the 5'-UTR of transcripts [111, 115]. In fact S1 has been suggested to be the first molecule to interact with the nascent RNA. The physiological role of this protein is not completely understood. It has been proposed that it could have an RNA unwinding activity that resolves structures in the 5'-UTR allowing the interaction with the 30S subunit [116]. S1 is dispensable for the translation of leaderless mRNAs, which are actually lacking in *E. coli* but are expressed by some *E. coli* bacteriophages. The association of S1 protein to the 30S subunit depends on the S2 protein that is needed to load S1 on the ribosome. It was demonstrated that in a S2 thermosensitive strain, at non- permissive temperature, ribosomes devoid of both S1 and S2 are generated. These ribosomes are able to translate leaderless but not canonical mRNAs [117]. Recently it has been shown that S1 interacts with S2 *via* the N-terminal domain [114]. In some Gram positive bacteria, such as *Mycobacterium tuberculosis*, natural leaderless mRNAs are common. Curiously, S1 protein is quite divergent in these bacteria and conserves only four out of the six S1 domains found in Gram-negative S1. Additional roles of S1 in other translation steps have never been directly addressed. In particular, it is not known whether S1 remains associated with the translating ribosome or if it is released after completion of translation initiation.

S1 has been involved in other phenomena besides translation. For instance, S1 can bind RNA polymerase; it has been claimed that S1 contribute in enhancing transcription from a number of

promoters and in the recycling of RNA polymerase [118, 119]. S1 has been involved in the interaction with PNPase and RNase E. Moreover it has been found associated to the poly(A) tail of mRNA [120, 121]. The interaction with factors involved in mRNA degradation suggests that S1 could play a role in the interplay between mRNA decay and translation. This was the object of my research, illustrated in the attached article 1 (Delvillani, F., Papiani, G., Dehò, G., and Briani, F. (2011). S1 ribosomal protein and the interplay between translation and mRNA decay. Nucl. Acids Res. 39(17):7702-15). We analysed the effect of altered S1 expression on the decay of model *cspE* and *rpsO* mRNAs and their leaderless variants. We demonstrated that S1 over-expression prevents mRNA-ribosome interaction, suggesting that S1 may have the paradoxical role of repressing translation when not assembled in the 30S particle. In fact, by cell extracts fractionation and western blotting of selected fractions, we showed that S1 over-expression leads to an increase of free S1. Part of the mRNA co-localizes with free S1 in the top fractions of the gradients. Interestingly, although translation is inhibited by S1 over-expression, the mRNA associated with free S1 is very stable. We could demonstrate that this is due to protection of the transcripts from RNase E-mediated degradation, whereas another decay pathway specific for *cspE* leaderless mRNA is not affected by S1. We speculate that S1 protection could preserve part of cellular mRNA from degradation during translational stress conditions. Moreover, our results suggest that ribosome devoid of S1 may exist in *E. coli* also in normal growth conditions. These ribosomes can translate only leaderless (and not leadered) mRNAs. It would be interesting to analyse whether the formation of such ribosomal particles can be modulated in response to external stimuli or growth phase.

RNA-based mechanisms of post-transcriptional regulation

Regulation by small non coding RNAs

Bacterial small RNAs (sRNAs) are a class of short regulatory RNAs usually ranging from 50 to 300 nt in length. They regulate the most disparate functions, from virulence to response to different physiological stresses. In Enterobacteria, many sRNAs regulate the expression of porins or other membrane proteins. Although some abundant cellular sRNAs such as 6S sRNA were early observed as specific bands in stained polyacrylamide gel, their role in cellular processes was not appreciated. The first evidence of a regulatory role of sRNAs dates back to the early ’80s, when the 108 nucleotide long RNAI was found to prevent ColE1 plasmid replication [122, 123]. The first identified chromosomally encoded sRNA was MicF, a translation inhibitor of the major outer membrane porin OmpF [124]. New sRNAs were then discovered through bioinformatics analysis aimed at identifying orphan promoters and terminators in the intergenic regions of sequenced genomes (reviewed by [125]). In the last years, transcriptomic approaches have been successfully applied to the discovery of new

sRNAs. Microarray and RNA deep sequencing have revealed the presence of hundreds of putative sRNAs in various Bacteria.

sRNAs can be divided in two classes: *cis* and *trans*-acting sRNAs. The first are transcribed in the same locus of their target genes, but from the opposite strand. sRNAs belonging to this class interact with their target RNAs by perfect base-pairing. *Trans* acting sRNAs are transcribed from a different locus that can be far from their target genes loci. They usually base-pair with imperfect complementarity to the target mRNA(s). The interaction between the sRNA and the mRNA target occurs *via* an initial contact involving few nucleotides in single stranded regions (kissing reaction). After this first interaction, other base-pairs can form that usually require secondary structures changes in both the sRNA and the mRNA.

In Gram-negative bacteria the interaction between trans-encoded sRNAs and targets is usually assisted by Hfq. This protein was initially discovered as the host factor needed for replication of the bacteriophage Qβ of *E. coli*[126]. Hfq forms a homohexamer that generates a central channel. Each protomer is formed by one α-helix and five β-strands. The ring-like architecture of Hfq exposes two faces as possible interactors for nucleic acids. The 'proximal face' is the surface that exposes the α-helices, whereas the "distal face" is at the opposite side. Hfq interacts with both sRNA and mRNA target. This interaction promotes and stabilizes the pairing [95].

sRNAs have different mode of action. In Enterobacteria, where they have been almost exclusively studied, they mostly act as translation repressors, by binding close to the TIR and inhibiting ribosome association and mRNA translation. An example of such mechanism is illustrated by regulation of *galK* gene by sRNA Spot42. Spot42 binds to the Shine-Dalgarno of *galK* mRNA preventing ribosome access and thereby inhibiting translation [127]. However, positive translation regulation has also been discovered. For example, the *rpoS* mRNA folds into a secondary structure that, by occluding the TIR, strongly inhibits translation. Three Hfq-dependent sRNAs (DsrA, RprA, and ArcZ) enhances *rpoS* translation. They all act in a similar way by pairing with the 5'-UTR of *rpoS* mRNA. This interaction opens the hairpin and allows mRNA translation by freeing the ribosome-binding site [128]. In some cases, sRNA-mRNA pairing can lead to degradation of the mRNA target. For example, the sRNA RyhB interacts with and destabilizes the *iscRSUA* mRNA [129].

Some sRNAs have multiple target RNAs. Spot42 directly regulates more than 10 operons involved in multiple aspects of cellular metabolism including the uptake and catabolism of diverse carbon sources [130]. Other sRNAs have a protein as target. Usually, the interaction with the sRNA inhibits the target protein activity. For instance, the sRNA CsrB has 18 binding sites for CsrA protein. CsrB binding inhibits CsrA activity by protein sequestering [131]. Another example of a sRNA acting through protein sequestering is the 6S RNA. This sRNA is very abundant in stationary

phase and is endowed with a secondary structure that mimics a DNA open complex. σ^{70}-RNA polymerase bind to 6S RNA [132]. Interestingly, the 6S RNA can be used as a template by RNA pol during outgrowth from stationary phase. This causes the dissociation of 6S-RNA pol complex [133].

Approaches to the identification of sRNA targets

Finding the target genes of *trans*-acting sRNAs represents a challenging task. The approaches that have been employed to this aim have intrinsic limitations, as they do not distinguish between direct and secondary targets, or are still scarcely reliable. Bioinformatic algorithms to predict sRNA targets evaluate the length of sRNA-mRNA base-pairing, the number of mismatches, base-pairing conservation rate among different species and its position (i.e. proximity to TIR region). Moreover, they consider the free energy of the pairing reaction i.e. the energetic balance between the costs of denaturing secondary structures in the sRNA and its target *vs.* the energy gain of the pairing between the two molecules. Bioinformatic tools usually generate long lists of putative targets with many false positives. Thus, the application of experimental methods is required to fish the actual target(s) among false positives. Moreover, actual targets are sometimes overlooked by bioinfomatic analysis.

Biochemical fishing of the target using the sRNA of interest as a bait has been documented only once. Douchin et al. applied this technique to fish the targets of the sRNA RseX. They synthetized *in vitro* a biotinylated RseX that was bound to streptavidin-coated magnetic beads. The beads were incubated with total *E. coli* RNA and the "captured" RNA was eluted and retro-transcribed to cDNA. Microarray analysis using this DNA as a probe revealed two stronger spots corresponding to *ompA* and *ompC* genes, which were then confirmed to be RseX actual targets [134].

Global proteomic analysis using 2D-gel has been used to detect differential protein expression relative to wild type by mutants with altered sRNA expression. In few lucky cases, it has been possible to find the target gene(s) of a given sRNA with a mono-dimensional protein gel. An example is the identification of the target genes of the *Agrobacterium tumefaciens* sRNA AbcR1/2 [135].

The approach that has been more frequently used to get sRNA targets is the comparative transcriptomic analysis (either by microarray or by RNA-Seq) in mutants with altered sRNA expression *vs*. the wt strain. An inherent limitation of this approach is that inhibition/enhancement of translation triggered by many sRNAs do not necessarily alter the stability of the mRNA target. Thus, as the amount of the mRNA target might not change, *bona fide* targets could be overlooked by this analysis. Moreover, both transcriptomic and proteomic approaches identify as potential targets all genes whose expression is affected by sRNA amount modulation, which are not

necessarily its direct targets. For instance, if a given small RNA regulates the expression of a global regulator, it will, indirectly, alter the expression of many genes.

P. aeruginosa sRNAs

Most studies on sRNAs have been performed in the closely related enterobacteria *E. coli* and *S. enterica*; in both these systems, about one hundred sRNAs have been found in different screenings. In comparison, when I started my Ph.D. training, the knowledge of sRNA complement of *P. aeruginosa* was quite poor. 22 sRNA genes were annotated in the Pseudomonas genome database (www.Pseudomonas.com). Only for six of them a function was suggested [136-140]. Moreover, experimental sRNA search had been performed only on PAO1 reference strain grown in standard laboratory conditions.

To discover new *P. aeruginosa* sRNAs, in collaboration with Prof. G. Bertoni group, we applied sRNA-Seq analysis of low molecular weight RNA extracted from two *P. aeuruginosa* strains (PAO1 and PA14) and validated by Northern Blot their expression. The two strains differ in pathogenicity degree, as PA14 is more virulent and contains two pathogenicity islands that are absent in PAO1. The rationale was that by comparing the sRNA complement expressed by the two strains, we could identify sRNAs potentially involved in virulence. In fact, in other pathogenic microbes, several sRNAs seem to be involved in the adaptation to the host environment [141]. Indeed, notwithstanding the function of only a very small number of *Pseudomonas* sRNAs is currently known, some of them have been implicated in virulence genes expression control (RsmY, Z) or in functions linked to virulence as quorum sensing (PhrS) [138, 142, 143] and other functions important for survival in the infected host, such as iron uptake and storage (PrrF1) [139].

Our work (attached file 2: Ferrara, S., Brugnoli, M., De Bonis, A., Righetti, F., Delvillani, F., Dehò, G., Horner, D., Briani, F., and Bertoni, G. (2012) Comparative Profiling of *Pseudomonas aeruginosa* Strains Reveals Differential Expression of Novel Unique and Conserved Small RNAs. PlosOne 7(5):e36553) identified more than 150 novel candidate sRNAs. We confirmed by Northern blotting the expression of 52 new sRNAs, substantially increasing the number of known sRNAs expressed by *P. aeruginosa*. Interestingly, a relevant number of new sRNAs were strain-specific or showed strain-specific expression, strongly suggesting that they could be involved in determining strain-characteristic phenotypic traits such as virulence degree. The research of the mRNA targets of a PA14-specific sRNA (SPA0021) is currently carried in our lab by the Tet-Trap genetic approach (see below).

Translation modulation by cis-acting structures

Riboswitches are a class of mRNA secondary structures that modulate the cognate (in *cis*) mRNA function in response to physical or chemical signals such as pH variation or metabolite binding. In particular, riboswitches can respond to several metabolites such as coenzymes, metal ions, tRNA or aminoacids [144, 145]. The environmental signals lead to changes in the riboswitch secondary structure that, at least in Gram negative bacteria, mostly affect mRNA translation (Fig. 5a). In general, in the presence of the ligand, a conformational switch occurs that can in turn release or sequester the ribosome binding site (RBS), thus modulating mRNA translation. For example vitamin B1 (thiamin) regulates the operons involved in its biosynthetic pathway [146]. These operons contain a region (thi-box) in the 5'-UTR that folds into a complex secondary structure. In the absence of the ligand (thiamin pyrophosphate, TPP) the thi-box relieves the intrinsic inhibitory effect of a structure involving the RBS. Conversely, when TPP is present, it binds the thi-box and the inhibitory structure can form [147, 148]. Some riboswitches may also regulate mRNA decay [149]. In Firmicutes, most riboswitches act by controlling premature transcription termination, as the ligand induces the formation of a Rho-independent terminator [150].

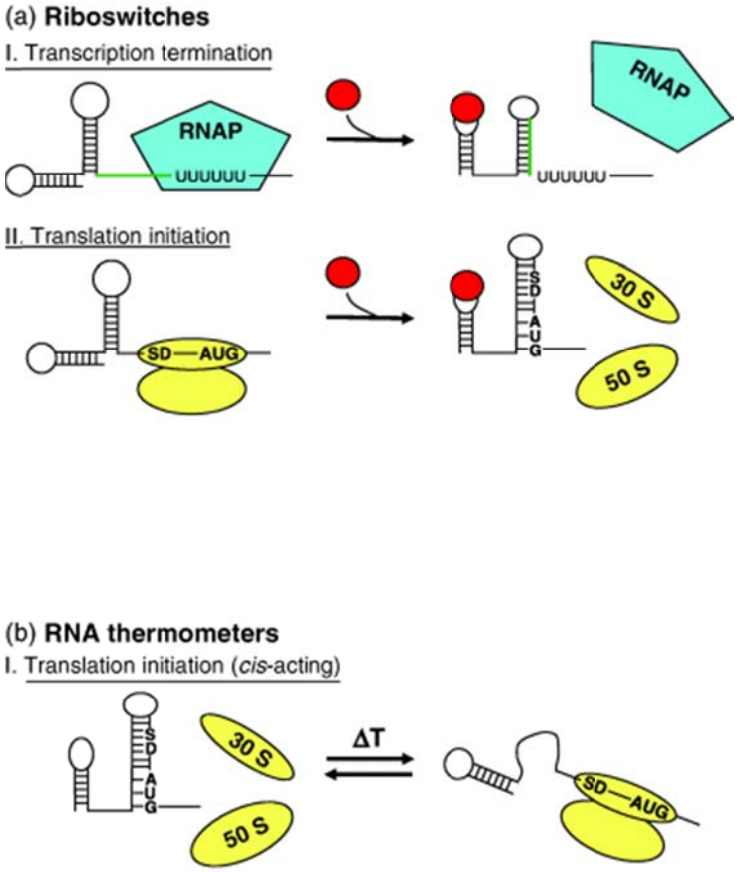


Figure 5. Post-transcriptional regulation by in *cis* elements. (a) Riboswitches. I) Transcription termination: the presence of a metabolite induces the formation of the Rho-independent hairpin that causes the release of RNA pol. II) Translation initiation: the metabolite alters mRNA secondary structure and this sequesters the Shine Dalgarno sequence. **(b)** RNA thermometers. Left: the translation initiation region is engaged in a secondary structure. Right: Temperature variation changes the secondary structure and releases the translation initiation region. From Narberhaus [151]

A particular class of in *cis* regulatory RNAs is represented by RNA thermometers (RNATs). They are exploited to monitor temperature changes in the surrounding environment. The thermosensor regulatory elements usually lie in the 5'-UTR of the genes. In most cases, at low temperature the translation initiation region is occluded by the RNAT secondary structure, which prevents ribosome binding and translation. As temperature increases, the RNAT structure is melted allowing 30S association (Fig. 5b).

In few cases, RNA thermometers repress translation at high temperature. They can fold into two different secondary structures in response to temperature changes that leave the TIR region accessible only at the lower temperature. The first RNAT identified belongs to this class and is

involved in the lysis-lysogeny decision of bacteriophage λ. λ cIII mRNA is present in two conformations in equilibrium between each other. The formation of the two structures depends on temperature and Mg^{2+} concentration; only one is accessible to ribosome and translated [152]. CIII is an alternative substrate for FtsH protease that degrades CII, a central protein in lysogenic state establishment [153, 154]. If CIII is expressed, CII protein is not degraded and promotes lysogeny. At high temperature (45° C), a stable structure occludes the TIR of CIII mRNA, thus preventing translation. CII protein is degraded by FtsH, its level decreases and lytic cycle occurs. This could be a mechanism evolved by λ bacteriophage to escape heat damages [155].

Only two classes of RNATs sharing common structural themes have been defined so far, ROSE and FourU elements; in both cases, sequence conservation within each class is limited to very short stretches of 4-5 bases in the proximity of the Shine-Dalgarno (SD) region [156, 157]. The ROSE (repression of heat shock gene expression) motif was discovered in the 5'-UTR of heat-shock small protein (hsp) operon of *Bradyrhizobium japonicum*, in particular in the 5'-UTR of the *hspA* gene. Translational fusion of *hspA* with *lacZ* showed low expression at 30°C. The expression increased when a 5'-UTR fragment was deleted and it was shown that the sequence itself acted as a translation repressor element at low temperature. Although the sequence of ROSE elements is not conserved, they all share some features. In all ROSEs, the Shine Dalgarno sequence is involved in a stem structure with a bulged "G" residue in the opposite strand. This residue is very relevant, as its elimination makes the thermosensor turned off also at high temperature, probably because the deletion optimizes the base-pairing and stabilizes the stem. All ROSEs have other (two or three) stem-loops besides that occluding the SD and AUG. In ROSEs the first nucleotide at the 5'-end of the transcript is paired. Co-transcriptional pairing seems to be necessary to promote correct folding. This idea is supported by the fact that heat-denatured ROSE RNAs don't turn to their original conformation when incubated at low temperature [158].

The FourU element was discovered in the *agsA* gene of *Salmonella enterica*. It is structurally simpler than ROSE as it is constituted by two hairpins; the SD sequence is paired with four uridine residues in hairpin II [157].

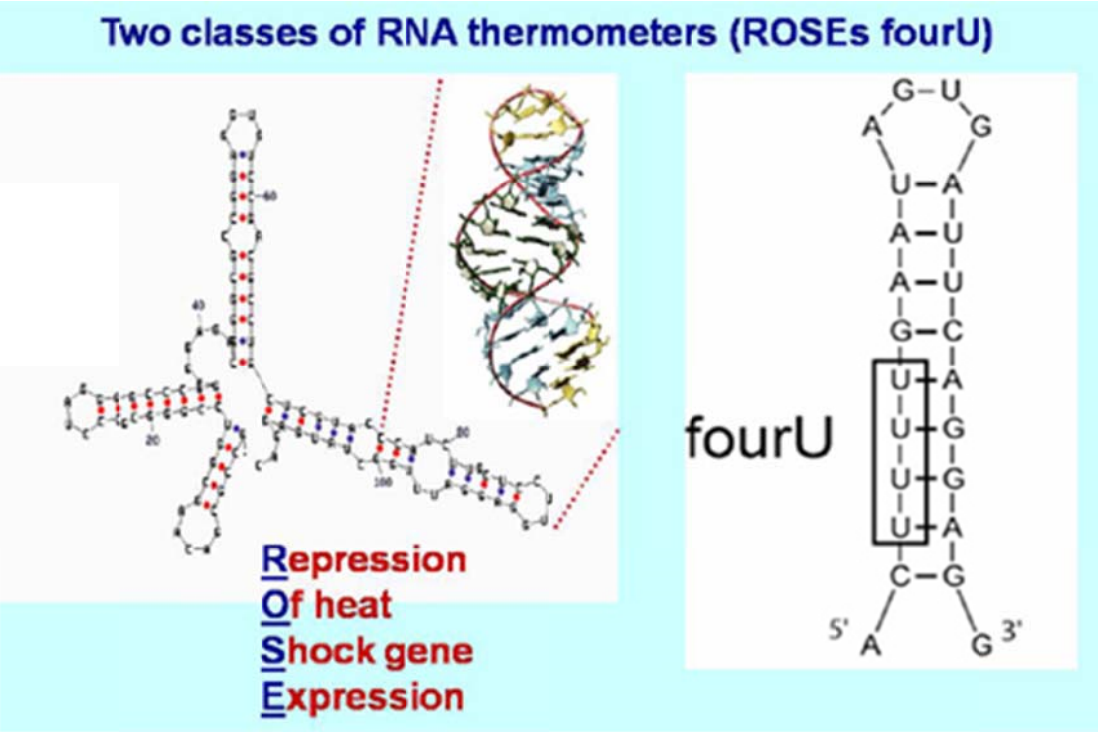


Figure 6. The two classes of conserved RNA thermometers. Left part: ROSE thermometer. Right: FourU thermometer. From: www.ruhr-uni-bochum.de/mikrobiologie/research.html.

Poor sequence and structure conservation hampers the bioinformatic search of RNATs. Moreover, RNATs structures may involve non-canonical base pairing (i.e. U-U), which are not considered by search algorithms [159]. On the other hand, the finding in the cyanobacterium *Synechocystis* sp. PCC 6803 of an RNAT constituted by a single stem-loop that modulates the expression of the small heat shock protein *hsp17* suggests that there might be many other as-yet undiscovered RNA thermometers in nature [160].

A strategy to identify new RNATs could be to focus on potentially temperature-controlled genes. In fact, not surprisingly, many RNA thermometers regulate the expression of heat-shock genes. However, this kind of strategy to the identification of RNATs would miss genes that are not obviously related to heat shock. In fact some pathogenic bacteria exploit the temperature upshift as a signal of the infection of warm-blooded hosts. In these bacteria, RNATs have been found to regulate the expression of virulence genes [155].

New P. aeruginosa RNATs identified by Tet-Trap

P. aeruginosa is a facultative pathogen, usually living in the soil at temperature well below 37 °C. It is thus conceivable that it can exploit the body temperature as a signal for the activation of virulence genes specifically required during the warm-blooded host infection. However, no RNA thermosensors have been found so far in this bacterium. In general, *P. aeruginosa* response to the

temperature upshift has been poorly explored. Only in 2012 a paper was published that compared the transcriptional profile of *P. aeruginosa* at 28° and 37°. The results of this analysis showed that 144 genes were overexpressed at 37° and 234 were repressed. Among genes up-regulated by temperature were most of those encoding T_3SS components and other genes that are known to play a role during infection [161].

In the attached manuscript (part III, Delvillani F. et al., A genetic approach to the identification of *Pseudomonas aeruginosa* RNA thermometers) we describe the identification of two putative RNATs modulating the expression of *dsbA* and *ptxS* genes. These RNATs were found by means of a genetic system, the Tet-Trap, which we developed, aimed at the identification of post-transcriptionally regulated genes. The Tet-Trap is based on Tn10 *tetA* gene regulation. In the absence of tetracycline, *tetA* transcription is prevented by TetR repressor binding to the operator *tetO*; tetracycline acts as allosteric inducer of TetR. A small peptide, TIP2, is able to mimic tetracycline effect on TetR as it leads to dissociation from *tetO*. TIP2 retains its properties also in end-fused chimeric polypeptides [162]. We have constructed *E. coli* strains to either select for or counter-select cells expressing TIP2-tagged proteins. By exploiting these strains, we could obtain a translational fusion library covering more than 60% of *P. aeruginosa* ORFs. The library was then probed for translation repression at low temperature. In a first screening, we analysed around 1200 clones and we found four genes that were post-transcriptionally regulated by temperature upshift from 28 to 37 °C. Thermoregulation of two of them (encoding the transcription factor PtxS and the periplasmic chaperone DsbA) was validated also with other reporter genes in *E. coli*. *ptxS* RNAT has been validated also in *P. aeruginosa*. The analysis of the other two putative RNATs identified in the screening is currently ongoing.

Interestingly, PtxS and DsbA were previously implicated in *P. aeruginosa* virulence [163, 164]. The investigation of the molecular mechanism and role in infection of temperature-dependent regulation of these two genes will be the subject of future research in our lab.

The Tet-Trap seems to be a reliable method to the identification of RNATs and it would be interesting to apply this analysis to other bacteria, such as *E. coli* pathogenic strains. Potentially, this approach could also be applied to the identification of genes regulated by riboswitches, provided that the ligand molecule is known and can be supplied to the cells. The application of Tet-Trap to the fishing of sRNA target genes is the object of current research in our lab.

References

1. Blattner, F.R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462 (1997).
2. Hu, P. et al. Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol* **7**, e96 (2009).
3. Croxen, M.A. & Finlay, B.B. Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* **8**, 26-38 (2010).
4. Macia, M.D. et al. Hypermutation is a key factor in development of multiple-antimicrobial resistance in *Pseudomonas aeruginosa* strains causing chronic lung infections. *Antimicrob Agents Chemother* **49**, 3382-6 (2005).
5. Coburn, G.A. & Mackie, G.A. Differential sensitivities of portions of the mRNA for ribosomal protein S20 to 3'-exonucleases dependent on oligoadenylation and RNA secondary structure. *J.Biol.Chem.* **271**, 15776-15781 (1996).
6. Cheng, Z.F. & Deutscher, M.P. Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. *J.Biol.Chem.* **277**, 21624-21629 (2002).
7. Blum, E., Carpousis, A.J. & Higgins, C.F. Polyadenylation promotes degradation of 3'-structured RNA by the *Escherichia coli* mRNA degradosome *in vitro J.Biol.Chem.* **274**, 4009-4016 (1999).
8. Deana, A. & Belasco, J.G. Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.* **19**, 2526-2533 (2005).
9. Braun, F., Hajnsdorf, E. & R,gnier, P. Polynucleotide phosphorylase is required for the rapid degradation of the RNase E-processed *rpsO* mRNA of *Escherichia coli* devoid of its 3' hairpin. *Mol.Microbiol.* **19**, 997-1005 (1996).
10. Marujo, P.E. et al. Inactivation of the decay pathway initiated at an internal site by RNase E promotes poly(A)-dependent degradation of the *rpsO* mRNA in *Escherichia coli*. *Mol.Microbiol.* **50**, 1283-1294 (2003).
11. Lehnik-Habrink, M., Lewis, R.J., Mader, U. & Stulke, J. RNA degradation in Bacillus subtilis: an interplay of essential endo- and exoribonucleases. *Mol.Microbiol.* (2012).
12. Condon, C. Maturation and degradation of RNA in bacteria. *Current Opinion in Microbiology* **10**, 271-278 (2007).
13. Apirion, D. & Lassar, A.B. A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. *J.Biol.Chem.* **253**, 1738-1742 (1978).
14. Kuwano, M. et al. Gene affecting longevity of messenger RNA: a mutant of *Escherichia coli* with altered mRNA stability. *Mol Gen Genet* **154**, 279-85 (1977).
15. Callaghan, A.J. et al. "Zn-link": a metal-sharing interface that organizes the quaternary structure and catalytic site of the endoribonuclease, RNase E. *Biochemistry* **44**, 4667-4675 (2005).
16. Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M. & Murzin, A.G. The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**, 235-242 (1997).
17. Cormack, R.S., Genereaux, J.L. & Mackie, G.A. RNase E activity is conferred by a single polypeptide: overexpression, purification, and properties of the *ams*/*rne*/*hmp1* gene product. *Proc.Natl.Acad.Sci.U.S.A* **90**, 9006-9010 (1993).
18. Vanzo, N.F. et al. Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* **12**, 2770-2781 (1998).
19. Kalman, M., Murphy, H. & Cashel, M. rhlB, a new *Escherichia coli* K-12 gene with an RNA helicase-like protein sequence motif, one of at least five such possible genes in a prokaryote. *New Biol.* **3**, 886-895 (1991).

20. Carpousis, A.J., Van Houwe, G., Ehretsmann, C. & Krisch, H.M. Copurification of E. coli RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* **76**, 889-900 (1994).
21. Py, B., Causton, H., Mudd, E.A. & Higgins, C.F. A protein complex mediating mRNA degradation in *Escherichia coli Mol.Microbiol.* **14**, 717-729 (1994).
22. Bernstein, J.A., Lin, P.H., Cohen, S.N. & Lin-Chao, S. Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc.Natl.Acad.Sci.U.S.A* **101**, 2758-2763 (2004).
23. Misra, T.K. & Apirion, D. RNase E, an RNA processing enzyme from *Escherichia coli*. *J.Biol.Chem.* **254**, 11154-11159 (1979).
24. Mackie, G.A., Genereaux, J.L. & Masterman, S.K. Modulation of the activity of RNase E *in vitro* by RNA sequences and secondary structures 5' to cleavage sites. *J.Biol.Chem.* **272**, 609-616 (1997).
25. Kim, K.S., Sim, S., Ko, J.H., Cho, B. & Lee, Y. Kinetic analysis of precursor M1 RNA molecules for exploring substrate specificity of the N-terminal catalytic half of RNase E. *J.Biochem.* **136**, 693-699 (2004).
26. Jiang, X. & Belasco, J.G. Catalytic activation of multimeric RNase E and RNase G by 5'-monophosphorylated RNA. *Proc.Natl.Acad.Sci.U.S.A* **101**, 9211-9216 (2004).
27. Mackie, G.A. RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* **11**, 45-57 (2013).
28. Murashko, O.N., Kaberdin, V.R. & Lin-Chao, S. Membrane binding of *Escherichia coli* RNase E catalytic domain stabilizes protein structure and increases RNA substrate affinity. *Proc Natl Acad Sci U S A* **109**, 7019-24 (2012).
29. Liou, G.G., Jane, W.N., Cohen, S.N., Lin, N.S. & Lin-Chao, S. RNA degradosomes exist in vivo in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc.Natl.Acad.Sci.U.S.A* **98**, 63-68 (2001).
30. Montero, L.P. et al. Spatial organization of the flow of genetic information in bacteria. *Nature* **466**, 77-81 (2010).
31. Bakshi, S., Siryaporn, A., Goulian, M. & Weisshaar, J.C. Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Mol Microbiol* **85**, 21-38 (2012).
32. Mudd, E.A. & Higgins, C.F. *Escherichia coli* endoribonuclease RNase E: autoregulation of expression and site-specific cleavage of mRNA. *Mol.Microbiol.* **9**, 557-568 (1993).
33. Singh, D. et al. Regulation of ribonuclease E activity by the L4 ribosomal protein of *Escherichia coli Proc.Natl.Acad.Sci.U.S.A* **106**, 864-869 (2009).
34. Gopel, Y., Papenfort, K., Reichenbach, B., Vogel, J. & Gorke, B. Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev* **27**, 552-64 (2013).
35. Li, Z., Pandit, S. & Deutscher, M.P. RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.* **18**, 2878-2885 (1999).
36. Umitsuki, G., Wachi, M., Takada, A., Hikichi, T. & Nagai, K. Involvement of RNase G in *in vivo* mRNA metabolism in *Escherichia coli*. *Genes Cells* **6**, 403-10 (2001).
37. Kaga, N., Umitsuki, G., Nagai, K. & Wachi, M. RNase G-dependent degradation of the eno mRNA encoding a glycolysis enzyme enolase in *Escherichia coli*. *Biosci Biotechnol Biochem* **66**, 2216-20 (2002).
38. Lee, K., Bernstein, J.A. & Cohen, S.N. RNase G complementation of rne null mutation identifies functional interrelationships with RNase E in *Escherichia coli*. *Mol.Microbiol.* **43**, 1445-1456 (2002).
39. Callahan, C., Neri-Cortes, D. & Deutscher, M.P. Purification and characterization of the tRNA-processing enzyme RNase BN. *J Biol Chem* **275**, 1030-4 (2000).

40. Perwez, T. & Kushner, S.R. RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Mol.Microbiol.* **60**, 723-737 (2006).
41. Apirion, D. & Watson, N. Mapping and characterization of a mutation in *Escherichia coli* that reduces the level of ribonuclease III specific for double-stranded ribonucleic acid. *J Bacteriol* **124**, 317-24 (1975).
42. Robertson, H.D. *Escherichia coli* ribonuclease III cleavage sites. *Cell* **30**, 669-672 (1982).
43. Zhang, K. & Nicholson, A.W. Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc.Natl.Acad.Sci.U.S.A* **94**, 13437-13441 (1997).
44. Court, D., Huang, T.F. & Oppenheim, A.B. Deletion analysis of the retroregulatory site for the lambda int gene. *J.Mol.Biol.* **166**, 233-240 (1983).
45. Portier, C., Dondon, L., Grunberg-Manago, M. & R,gnier, P. The first step in the functional inactivation of the *Escherichia coli* polynucleotide phosphorylase messenger is a ribonuclease III processing at the 5' end. *EMBO J.* **6**, 2165-2170 (1987).
46. Bardwell, J.C. et al. Autoregulation of RNase III operon by mRNA processing. *EMBO J.* **8**, 3401-3407 (1989).
47. Viegas, S.C., Silva, I.J., Saramago, M., Domingues, S. & Arraiano, C.M. Regulation of the small regulatory RNA MicA by ribonuclease III: a target-dependent pathway. *Nucleic Acids Res.* **39**, 2918-2930 (2011).
48. Cesareni, G., Helmer-Citterich, M. & Castagnoli, L. Control of ColE1 plasmid replication by antisense RNA. *Trends Genet* **7**, 230-5 (1991).
49. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-9 (2003).
50. Talbot, S.J. & Altman, S. Gel retardation analysis of the interaction between C5 protein and M1 RNA in the formation of the ribonuclease P holoenzyme from *Escherichia coli*. *Biochemistry* **33**, 1399-405 (1994).
51. Guerrier-Takada, C., Subramanian, A.R. & Cole, P.E. The activity of discrete fragments of ribosomal protein S1 in Q beta replicase function. *J.Biol.Chem.* **258**, 13649-13652 (1983).
52. Liu, J.M. et al. Experimental discovery of sRNAs in Vibrio cholerae by direct cloning, 5S/tRNA depletion and parallel sequencing LIU2009. *Nucleic Acids Res.* **37**, e46 (2009).
53. Li, Y. & Altman, S. Polarity effects in the lactose operon of *Escherichia coli*. *J Mol Biol* **339**, 31-9 (2004).
54. Alifano, P. et al. Ribonuclease E provides substrates for ribonuclease P-dependent processing of a polycistronic mRNA. *Genes Dev* **8**, 3021-31 (1994).
55. Andrade, J.M., Pobre, V., Silva, I.J., Domingues, S. & Arraiano, C.M. The role of 3'-5' exoribonucleases in RNA degradation. *Prog Mol Biol Transl Sci* **85**, 187-229 (2009).
56. Zuo, Y. & Deutscher, M.P. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* **29**, 1017-1026 (2001).
57. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell* **91**, 457-66 (1997).
58. Amblar, M., Barbas, A., Fialho, A.M. & Arraiano, C.M. Characterization of the functional domains of *Escherichia coli* RNase II. *J Mol Biol* **360**, 921-33 (2006).
59. 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).
60. Deutscher, M.P. & Reuven, N.B. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis Proc.Natl.Acad.Sci.U.S.A* **88**, 3277-3280 (1991).
61. Cao, G.J., Kalapos, M.P. & Sarkar, N. Polyadenylated mRNA in *Escherichia coli* : modulation of poly(A) RNA levels by polynucleotide phosphorylase and ribonuclease II. *Biochimie* **79**, 211-220 (1997).
62. Marujo, P.E. et al. RNase II removes the oligo(A) tails that destabilize the *rpsO* mRNA of *Escherichia coli*. *RNA.* **6**, 1185-1193 (2000).

63. Awano, N. et al. *Escherichia coli* RNase R has dual activities, helicase and RNase. *J Bacteriol* **192**, 1344-52 (2010).
64. Vincent, H.A. & Deutscher, M.P. Substrate recognition and catalysis by the exoribonuclease RNase R. *J Biol Chem* **281**, 29769-75 (2006).
65. Andrade, J.M., Cairrao, F. & Arraiano, C.M. RNase R affects gene expression in stationary phase: regulation of *ompA*. *Molecular Microbiology* **60**, 219-228 (2006).
66. Cairrao, F., Cruz, A., Mori, H. & Arraiano, C.M. Cold shock induction of RNase R and its role in the maturation of the quality control mediator *SsrA/tmRNA*. *Mol.Microbiol.* **50**, 1349-1360 (2003).
67. Mohanty, B.K. & Kushner, S.R. Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in *Escherichia coli* *Proc.Natl.Acad.Sci.U.S.A* **97**, 11966-11971 (2000).
68. Thang, D.C. [Purification and characterization of the polynucleotide phosphorylase of Azotobacter vinelandii]. *Bull.Soc.Chim.Biol.(Paris)* **49**, 1773-1783 (1967).
69. Klee, C.B. & Singer, M.F. The processive degradation of individual polyribonucleotide chains. II. Micrococcus lysodeikticus polynucleotide phosphorylase. *J.Biol.Chem.* **243**, 923-927 (1968).
70. Chen, H.W. et al. Mammalian polynucleotide phosphorylase is an intermembrane space RNase that maintains mitochondrial homeostasis. *Molecular and Cellular Biology* **26**, 8475-8487 (2006).
71. Vedrenne, V. et al. Mutation in PNPT1, which encodes a polyribonucleotide nucleotidyltransferase, impairs RNA import into mitochondria and causes respiratory-chain deficiency. *Am J Hum Genet* **91**, 912-8 (2012).
72. Das, S.K. et al. Human polynucleotide phosphorylase (hPNPase(old-35)): an evolutionary conserved gene with an expanding repertoire of RNA degradation functions. *Oncogene* **30**, 1733-1743 (2011).
73. Symmons, M.F., Jones, G.H. & Luisi, B.F. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure.Fold.Des* **8**, 1215-1226 (2000).
74. Shi, Z., Yang, W.Z., Lin-Chao, S., Chak, K.F. & Yuan, H.S. Crystal structure of *Escherichia coli* PNPase: central channel residues are involved in processive RNA degradation. *RNA.* **14**, 2361-2371 (2008).
75. Stickney, L.M., Hankins, J.S., Miao, X. & Mackie, G.A. Function of the conserved S1 and KH domains in polynucleotide phosphorylase. *J.Bacteriol.* **187**, 7214-7221 (2005).
76. Matus-Ortega, M.E. et al. The KH and S1 domains of *Escherichia coli* polynucleotide phosphorylase are necessary for autoregulation and growth at low temperature. *Biochimica et Biophysica Acta-Gene Structure and Expression* **1769**, 194-203 (2007).
77. Briani, F. et al. Genetic analysis of polynucleotide phosphorylase structure and functions. *Biochimie* **89**, 145-157 (2007).
78. Liou, G.G., Chang, H.Y., Lin, C.S. & Lin-Chao, S. DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E. *J.Biol.Chem.* **277**, 41157-41162 (2002).
79. Luttinger, A., Hahn, J. & Dubnau, D. Polynucleotide phosphorylase is necessary for competence development in *Bacillus subtilis Mol.Microbiol.* **19**, 343-356 (1996).
80. Zangrossi, S. et al. Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in *Escherichia coli Mol.Microbiol.* **36**, 1470-1480 (2000).
81. Siculella, L. et al. Guanosine 5'-diphosphate 3'-diphosphate (ppGpp) as a negative modulator of polynucleotide phosphorylase activity in a 'rare' actinomycete. *Mol.Microbiol.* **77**, 716-729 (2010).

82. Nurmohamed, S. et al. Polynucleotide phosphorylase activity may be modulated by metabolites in *Escherichia coli J.Biol.Chem.* **286**, 14315-14323 (2011).
83. Del Favero, M. et al. Regulation of *Escherichia coli* polynucleotide phosphorylase by ATP. *J.Biol.Chem.* **283**, 27355-27359 (2008).
84. Tuckerman, J.R., Gonzalez, G. & Gilles-Gonzalez, M.A. Cyclic di-GMP activation of polynucleotide phosphorylase signal-dependent RNA processing. *J.Mol.Biol.* **407**, 633-639 (2011).
85. Briani, F. et al. Polynucleotide phosphorylase hinders mRNA degradation upon ribosomal protein S1 overexpression in *Escherichia coli RNA.* **14**, 2417-2429 (2008).
86. De Lay, N. & Gottesman, S. Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli RNA.* **17**, 1172-1189 (2011).
87. Cardenas, P.P. et al. *Bacillus subtilis* polynucleotide phosphorylase 3'-to-5' DNase activity is involved in DNA repair. *Nucleic Acids Res.* **37**, 4157-4169 (2009).
88. Cardenas, P.P. et al. Polynucleotide phosphorylase exonuclease and polymerase activities on single-stranded DNA ends are modulated by RecN, SsbA and RecA proteins. *Nucleic Acids Res.* **39**, 9250-9261 (2011).
89. Jarrige, A.C., Mathy, N. & Portier, C. PNPase autocontrols its expression by degrading a double-stranded structure in the pnp mRNA leader. *EMBO J.* **20**, 6845-6855 (2001).
90. Carzaniga, T. et al. Autogenous regulation of *Escherichia coli* polynucleotide phosphorylase expression revisited. *J.Bacteriol.* **191**, 1738-1748 (2009).
91. Ghosh, S. & Deutscher, M.P. Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc.Natl.Acad.Sci.U.S.A* **96**, 4372-4377 (1999).
92. Coburn, G.A. & Mackie, G.A. Degradation of mRNA in *Escherichia coli*: an old problem with some new twists. *Prog.Nucleic Acid Res.Mol.Biol.* **62**, 55-108 (1999).
93. Mohanty, B.K. & Kushner, S.R. The majority of *Escherichia coli* mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucleic Acids Research* **34**, 5695-5704 (2006).
94. Mohanty, B.K. & Kushner, S.R. Deregulation of poly(A) polymerase I in *Escherichia coli* inhibits protein synthesis and leads to cell death. *Nucleic Acids Res* **41**, 1757-66 (2013).
95. Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V.R. & Blasi, U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA.* **9**, 1308-1314 (2003).
96. Hajnsdorf, E. & Regnier, P. Host factor Hfq of *Escherichia coli* stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proc Natl Acad Sci U S A* **97**, 1501-5 (2000).
97. Deana, A., Celesnik, H. & Belasco, J.G. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**, 355-358 (2008).
98. Luciano, D.J. et al. Differential control of the rate of 5'-end-dependent mRNA degradation in *Escherichia coli*. *J Bacteriol* **194**, 6233-9 (2012).
99. Lindahl, L. & Zengel, J.M. Ribosomal genes in *Escherichia coli*. *Annu Rev Genet* **20**, 297-326 (1986).
100. Squires, C.L., Greenblatt, J., Li, J., Condon, C. & Squires, C.L. Ribosomal RNA antitermination *in vitro*: requirement for Nus factors and one or more unidentified cellular components. *Proc Natl Acad Sci U S A* **90**, 970-4 (1993).
101. Zengel, J.M. & Lindahl, L. *Escherichia coli* ribosomal protein L4 stimulates transcription termination at a specific site in the leader of the S10 operon independent of L4-mediated inhibition of translation. *J.Mol.Biol.* **213**, 67-78 (1990).
102. Schlax, P.J. & Worhunsky, D.J. Translational repression mechanisms in prokaryotes. *Mol Microbiol* **48**, 1157-69 (2003).
103. Boni, I.V., Isaeva, D.M., Musychenko, M.L. & Tzareva, N.V. Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res.* **19**, 155-162 (1991).

104. Schmeing, T.M. & Ramakrishnan, V. What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**, 1234-42 (2009).
105. Kaberdina, A.C., Szaflarski, W., Nierhaus, K.H. & Moll, I. An unexpected type of ribosomes induced by kasugamycin: a look into ancestral times of protein synthesis? *Mol.Cell* **33**, 227-236 (2009).
106. Okuyama, A., Machiyama, N., Kinoshita, T. & Tanaka, N. Inhibition by kasugamycin of initiation complex formation on 30S ribosomes. *Biochem Biophys Res Commun* **43**, 196-9 (1971).
107. Amitai, S., Kolodkin-Gal, I., Hananya-Meltabashi, M., Sacher, A. & Engelberg-Kulka, H. *Escherichia coli* MazF leads to the simultaneous selective synthesis of both "death proteins" and "survival proteins". *PLoS Genet* **5**, e1000390 (2009).
108. Vesper, O. et al. Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell* **147**, 147-57 (2011).
109. Dreyfus, M. Killer and protective ribosomes. *Prog.Mol.Biol.Transl.Sci.* **85**, 423-466 (2009).
110. Kimura, M., Foulaki, K., Subramanian, A.R. & Wittmann-Liebold, B. Primary structure of *Escherichia coli* ribosomal protein S1 and features of its functional domains. *Eur.J.Biochem.* **123**, 37-53 (1982).
111. Subramanian, A.R. Structure and functions of ribosomal protein S1. *Prog.Nucleic Acid Res.Mol.Biol.* **28**, 101-142 (1983).
112. Murzin, A.G. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861-867 (1993).
113. Theobald, D.L., Mitton-Fry, R.M. & Wuttke, D.S. Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32**, 115-33 (2003).
114. Byrgazov, K., Manoharadas, S., Kaberdina, A.C., Vesper, O. & Moll, I. Direct interaction of the N-terminal domain of ribosomal protein S1 with protein S2 in *Escherichia coli*. *PLoS One* **7**, e32702 (2012).
115. Komarova, A.V., Tchufistova, L.S., Supina, E.V. & Boni, I.V. Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. *RNA.* **8**, 1137-1147 (2002).
116. Bear, D.G. et al. Alteration of polynucleotide secondary structure by ribosomal protein S1. *Proc.Natl.Acad.Sci.U.S.A* **73**, 1824-1828 (1976).
117. Moll, I., Grill, S., Grundling, A. & Blasi, U. Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in *Escherichia coli*. *Mol.Microbiol.* **44**, 1387-1396 (2002).
118. Sukhodolets, M.V. & Garges, S. Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42**, 8022-8034 (2003).
119. Sukhodolets, M.V., Garges, S. & Adhya, S. Ribosomal protein S1 promotes transcriptional cycling. *RNA* **12**, 1505-1513 (2006).
120. Kalapos, M.P., Paulus, H. & Sarkar, N. Identification of ribosomal protein S1 as a poly(A) binding protein in *Escherichia coli Biochimie* **79**, 493-502 (1997).
121. Feng, Y., Huang, H., Liao, J. & Cohen, S.N. *Escherichia coli* poly(A)-binding proteins that interact with components of degradosomes or impede RNA decay mediated by polynucleotide phosphorylase and RNase E. *J.Biol.Chem.* **276**, 31651-31656 (2001).
122. Stougaard, P., Molin, S. & Nordstrom, K. RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc.Natl.Acad.Sci.U.S.A* **78**, 6008-6012 (1981).
123. Tomizawa, J., Itoh, T., Selzer, G. & Som, T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc Natl Acad Sci U S A* **78**, 1421-5 (1981).
124. Mizuno, T. [Regulation of gene expression by a small RNA transcript (micRNA): osmoregulation in E. coli]. *Seikagaku* **56**, 113-9 (1984).
125. Livny, J. Efficient annotation of bacterial genomes for small, noncoding RNAs using the integrative computational tool sRNAPredict2. *Methods Mol Biol* **395**, 475-88 (2007).

126. August, J.T. et al. Phage-specific and host proteins in the replication of bacteriophage RNA. *Fed Proc* **29**, 1170-5 (1970).
127. Moller, T., Franch, T., Udesen, C., Gerdes, K. & Valentin-Hansen, P. Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon. *Genes Dev* **16**, 1696-706 (2002).
128. Battesti, A., Majdalani, N. & Gottesman, S. The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* **65**, 189-213 (2011).
129. Desnoyers, G., Morissette, A., Prevost, K. & Masse, E. Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. *EMBO J.* **28**, 1551-1561 (2009).
130. Beisel, C.L. & Storz, G. Discriminating tastes: Physiological contributions of the Hfq-binding small RNA Spot 42 to catabolite repression. *RNA.Biol.* **8** (2011).
131. Babitzke, P. & Romeo, T. CsrB sRNA family: sequestration of RNA-binding regulatory proteins BABITZKE2007. *Curr.Opin.Microbiol.* **10**, 156-163 (2007).
132. Wassarman, K.M. 6S RNA: a small RNA regulator of transcription. *Curr Opin Microbiol* **10**, 164-8 (2007).
133. Wassarman, K.M. & Saecker, R.M. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. *Science* **314**, 1601-3 (2006).
134. Douchin, V., Bohn, C. & Bouloc, P. Down-regulation of porins by a small RNA bypasses the essentiality of the regulated intramembrane proteolysis protease RseP in *Escherichia coli*. *J Biol Chem* **281**, 12253-9 (2006).
135. Wilms, I., Voss, B., Hess, W.R., Leichert, L.I. & Narberhaus, F. Small RNA-mediated control of the Agrobacterium tumefaciens GABA binding protein. *Mol.Microbiol.* **80**, 492-506 (2011).
136. Sonnleitner, E. et al. Detection of small RNAs in *Pseudomonas aeruginosa* by RNomics and structure-based bioinformatic tools SONNLEITNER2008. *Microbiology* **154**, 3175-3187 (2008).
137. Livny, J., Brencic, A., Lory, S. & Waldor, M.K. Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2 LIVNY2006. *Nucleic Acids Res.* **34**, 3484-3493 (2006).
138. Heurlier, K. et al. Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. *J.Bacteriol.* **186**, 2936-2945 (2004).
139. Wilderman, P.J. et al. Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis WILDERMAN2004. *Proc.Natl.Acad.Sci.U.S.A* **101**, 9792-9797 (2004).
140. Toschka, H.Y., Struck, J.C. & Erdmann, V.A. The 4.5S RNA gene from *Pseudomonas aeruginosa*. *Nucleic Acids Res* **17**, 31-6 (1989).
141. Papenfort, K. & Vogel, J. Regulatory RNA in bacterial pathogens. *Cell Host Microbe* **8**, 116-27 (2010).
142. Valverde, C., Heeb, S., Keel, C. & Haas, D. RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0. *Mol.Microbiol.* **50**, 1361-1379 (2003).
143. Sonnleitner, E. et al. The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol.Microbiol.* **80**, 868-885 (2011).
144. Henkin, T.M. RNA-dependent RNA switches in bacteria. *Methods Mol Biol* **540**, 207-14 (2009).
145. Spinelli, S.V., Pontel, L.B., Garcia Vescovi, E. & Soncini, F.C. Regulation of magnesium homeostasis in Salmonella: Mg(2+) targets the mgtA transcript for degradation by RNase E. *FEMS Microbiol Lett* **280**, 226-34 (2008).
146. Miranda-Rios, J. The THI-box riboswitch, or how RNA binds thiamin pyrophosphate. *Structure* **15**, 259-65 (2007).

147. Stormo, G.D. & Ji, Y. Do mRNAs act as direct sensors of small molecules to control their expression? *Proc Natl Acad Sci U S A* **98**, 9465-7 (2001).
148. Miranda-Rios, J., Navarro, M. & Soberon, M. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc Natl Acad Sci U S A* **98**, 9736-41 (2001).
149. Collins, J.A., Irnov, I., Baker, S. & Winkler, W.C. Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev* **21**, 3356-68 (2007).
150. Barrick, J.E. & Breaker, R.R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* **8**, R239 (2007).
151. Narberhaus, F., Waldminghaus, T. & Chowdhury, S. RNA thermometers. *FEMS Microbiol.Rev.* **30**, 3-16 (2006).
152. Altuvia, S., Kornitzer, D., Teff, D. & Oppenheim, A.B. Alternative mRNA structures of the *cIII* gene of bacteriophage lambda determine the rate of its translation initiation. *J Mol Biol* **210**, 265-80 (1989).
153. Herman, C., Thevenet, D., D'Ari, R. & Bouloc, P. The HflB protease of *Escherichia coli* degrades its inhibitor lambda cIII. *J Bacteriol* **179**, 358-63 (1997).
154. Shotland, Y. et al. Proteolysis of the phage lambda CII regulatory protein by FtsH (HflB) of *Escherichia coli*. *Mol Microbiol* **24**, 1303-10 (1997).
155. Narberhaus, F. Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. *RNA Biol* **7**, 84-9 (2010).
156. Waldminghaus, T., Fippinger, A., Alfsmann, J. & Narberhaus, F. RNA thermometers are common in alpha- and gamma-proteobacteria. *Biol Chem* **386**, 1279-86 (2005).
157. Waldminghaus, T., Heidrich, N., Brantl, S. & Narberhaus, F. FourU: a novel type of RNA thermometer in Salmonella. *Mol Microbiol* **65**, 413-24 (2007).
158. Chowdhury, S., Ragaz, C., Kreuger, E. & Narberhaus, F. Temperature-controlled structural alterations of an RNA thermometer. *J Biol Chem* **278**, 47915-21 (2003).
159. Chowdhury, S., Maris, C., Allain, F.H. & Narberhaus, F. Molecular basis for temperature sensing by an RNA thermometer. *EMBO J* **25**, 2487-97 (2006).
160. Kortmann, J., Sczodrok, S., Rinnenthal, J., Schwalbe, H. & Narberhaus, F. Translation on demand by a simple RNA-based thermosensor. *Nucleic Acids Res* **39**, 2855-68 (2011).
161. Wurtzel, O. et al. The single-nucleotide resolution transcriptome of *Pseudomonas aeruginosa* grown in body temperature. *PLoS Pathog* **8**, e1002945 (2012).
162. Goeke, D. et al. Short peptides act as inducers, anti-inducers and corepressors of tet repressor. *J.Mol.Biol.* **416**, 33-45 (2012).
163. Colmer, J.A. & Hamood, A.N. Characterization of ptxS, a *Pseudomonas aeruginosa* gene which interferes with the effect of the exotoxin A positive regulatory gene, ptxR. *Mol Gen Genet* **258**, 250-9 (1998).
164. Ha, U.H., Wang, Y. & Jin, S. DsbA of *Pseudomonas aeruginosa* is essential for multiple virulence factors. *Infect Immun* **71**, 1590-5 (2003).

PART II

Content:

Research article 1. Delvillani, F., Papiani, G., Dehò, G., and Briani, F. (2011). S1 ribosomal protein and the interplay between translation and mRNA decay. Nucl. Acids Res. 39(17):7702-15

Research article 2. Ferrara, S., Brugnoli, M., De Bonis, A., Righetti, F., Delvillani, F., Dehò, G., Horner, D., Briani, F., and Bertoni, G. (2012) Comparative Profiling of *Pseudomonas aeruginosa* Strains Reveals Differential Expression of Novel Unique and Conserved Small RNAs. PlosOne 7(5):e36553

Nucleic Acids Research Advance Access published June 17, 2011

Nucleic Acids Research, 2011, 1–14
doi:10.1093/nar/gkr417

S1 ribosomal protein and the interplay between translation and mRNA decay

Francesco Delvillani, Giulia Papiani, Gianni Dehò and Federica Briani*

Dipartimento di Scienze biomolecolari e Biotecnologie, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy

Received March 16, 2011; Revised May 2, 2011; Accepted May 6, 2011

ABSTRACT

S1 is an 'atypical' ribosomal protein weakly associated with the 30S subunit that has been implicated in translation, transcription and control of RNA stability. S1 is thought to participate in translation initiation complex formation by assisting 30S positioning in the translation initiation region, but little is known about its role in other RNA transactions. In this work, we have analysed *in vivo* the effects of different intracellular S1 concentrations, from depletion to overexpression, on translation, decay and intracellular distribution of leadered and leaderless messenger RNAs (mRNAs). We show that the *cspE* mRNA, like the *rpsO* transcript, may be cleaved by RNase E at multiple sites, whereas the leaderless *cspE* transcript may also be degraded via an alternative pathway by an unknown endonuclease. Upon S1 overexpression, RNase E-dependent decay of both *cspE* and *rpsO* mRNAs is suppressed and these transcripts are stabilized, whereas cleavage of leaderless *cspE* mRNA by the unidentified endonuclease is not affected. Overall, our data suggest that ribosome-unbound S1 may inhibit translation and that part of the *Escherichia coli* ribosomes may actually lack S1.

INTRODUCTION

Decades of research in the model organism *Escherichia coli* have provided a deep knowledge of cellular machineries involved in translation and messenger RNA (mRNA) degradation; however, how these two processes are interconnected at the molecular level is still poorly understood. It is commonly accepted that translation deeply affects mRNA decay, as mutations that prevent or reduce translation usually shorten mRNA half-life. However, a relatively low number of studies have directly addressed the interplay between translation and RNA degradation and a small repertoire of model mRNAs have been analysed in this respect so far (1,2).

Serendipitous observations by different laboratories suggest that the ribosomal protein S1 could be involved in the crosstalk between protein synthesis and RNA degradation. S1 is the largest ribosomal protein in the 30S subunit of *E. coli* ribosome and is the only ribosomal protein with documented high affinity for mRNA (3). The protein has also been identified as a poly(A) tail binding factor from *E. coli* cell extracts (4) and shown to interact with RNase E and PNPase, two of the main *E. coli* RNA degrading enzymes, in Far-Western assays (5). Moreover, altering S1 expression from overexpression to depletion has opposite effects on mRNA expression, since S1 excess seems to stabilize different *E. coli* mRNAs that become barely detectable upon S1 depletion (6).

S1 has been considered a translation factor rather than a 'real' ribosomal protein, given its weak and reversible association with ribosomes (7,8) and its stoichiometry of less than one copy per 30S subunit (9). However, dissociation of S1 from the 30S subunit after cell lysis has been considered by different groups an experimental artefact, thus questioning the stoichiometry of the protein in the ribosome and the real magnitude of the non-ribosomal S1 pool (10–12). As a matter of fact, S1 is one of the few ribosomal proteins whose role in translation has been specifically analysed. *In vivo,* S1 is essential for growth and is required for translation of bulk mRNA in *E. coli* (6,13); on the other hand, ribosomes depleted of S1 and S2 retain the ability of translating the naturally leaderless λ *cI* and Tn*1721 tetR* mRNAs (14). Recently, it has been shown that a 'minimal' ribosome, lacking several proteins of the 30S subunit, among which S1, is still proficient in leaderless mRNA translation (15). *In vitro*, S1 is required for the assembly of 30S initiation complex at internal ribosome entry sites [i.e. located in 5′-untranslated region (UTR)] of mRNAs (16). It has been proposed that the interaction

*To whom correspondence should be addressed. Tel: +39 02 5031 5033; Fax: +39 02 5031 5044; Email: federica.briani@unimi.it
Present address:
Giulia Papiani, CNR, Istituto di Neuroscienze, Via Vanvitelli 32, 20129 Milano, Italy

between S1 in the 30S and the mRNA 5′-UTR may be responsible for a first, rapid and reversible step in initiation complex formation, which will be followed by the establishment of specific interactions between the Shine–Dalgarno (SD) and 16S rRNA (17,18). However, S1 is dispensable for initiation complex formation on RNAs with a strong SD region (19) or on leaderless mRNAs. In fact, initiation of leaderless mRNA translation occurs *in vitro* through a non-conventional pathway by direct binding to the 70S ribosome (14,20,21). This 70S-dependent initiation pathway seems to be, at least *in vitro*, insensitive to the presence of S1, since it occurs with ribosomes devoid of both S1 and S2 (as a consequence of an *rpsBts* mutation) and with crosslinked 70S ribosomes still containing S1 and S2 (20). It has been reported that also *in vivo*, in conditions where 70S ribosomes become prevalent because of a mutation that impairs ribosome recycling factor activity, leaderless mRNAs are translated whereas translation of bulk mRNA ceases (20). It is not known whether the 70S particles that accumulate in the mutant still retain S1.

S1 binding sites on mRNA have been recognized as A/U- rich single-stranded regions usually located immediately upstream of the SD (22,23). Interestingly, those regions constitute RNase E cleavage sites in different mRNAs (1). Nevertheless, the insertion of AU-rich elements upstream of an SD sequence enhances translation and stabilizes mRNA, suggesting that ribosome assembly on the mRNA via S1 binding may prevent RNase E cutting (24). It remains to be established, however, whether *in vivo* S1 not bound to the ribosome may also interact with mRNA and regulate its decay.

We have previously shown that both S1 overexpression and depletion inhibit bacterial growth but have different outcomes on mRNA expression (6). We observed that upon S1 depletion, the amount of several mRNAs sharply decreased; conversely, the quantity of most mRNAs did not significantly change or increase in S1 over-expressing cells. However, upon S1 overexpression, all the assayed mRNAs became notably more stable than in S1 basal expression condition. Surprisingly, the exonuclease polynucleotide phosphorylase (PNPase) seemed to enhance S1 protective effect for most of the assayed mRNAs.

In this work, we have investigated the role of mRNA association with the ribosome and translation on S1-dependent modulation of mRNA stability. Our data suggest that S1 may specifically inhibit RNase E-dependent decay by hindering RNase E cleavage sites.

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains and plasmids are listed in Supplementary Table S1. *Escherichia coli* sequence coordinates are from NCBI Accession Number U00096.2. C-1a (25), C-5868 and C-5869 (26) have been previously described. C-5699 is a C-5698 derivative (6) in which the *cat* resistance cassette was excised by FLP-mediated recombination as described (27). C-5874 was obtained by P1 transduction of the Δ*cspE*::*kan* allele from JW0618 [Keio collection; (28)] into C-1a; the resistance cassette was then excised by FLP-mediated recombination. C-5899 and C-5901 were obtained by P1 transduction of the Δ*rng*::*kan* and Δ*elaC*::*kan* alleles from JW3216 and JW2263 strains [Keio collection; (28)], respectively, in C-5868.

pQE31-S1 and pREP4 were kindly provided by M. V. Sukhodolets. The recombinant S1 protein expressed from pQE31-S1 allele carries an N-terminal His6 tag (29). The other plasmids used in this work (see also Figure 2A) are derivatives of pGM385. This plasmid was obtained by cloning in the XbaI site of pGM743 vector (30,31) a 170-bp-long DNA fragment carrying the bacteriophage P4 Rho-dependent transcription terminator t_{imm} (GenBank Accession Number X51522: 8365-8209) (32). Coordinates of the *E. coli cspE* fragments cloned in pGM385 are reported in Supplementary Table S1. In plasmids pGM928 and pGM929, the HA epitope coding fragment (TACCCATACGACGTCCCAGACTACGCT) has been inserted in frame within the *cspE* coding region. Plasmids pGM396 and pGM398 carry the *rpsO* gene with or without the 5′-UTR, respectively. The *rpsO* region essential for the interaction with the ribosome (33) has been deleted in these plasmids and replaced with the in frame HA epitope DNA, flanked by NheI and SacI restriction sites. Plasmid pGM397 is a pGM396 derivative in which the *rpsO* region upstream of the HA has been replaced by an in frame phage λ DNA fragment carrying P_{RM} and the first 189 bp of *cI* open reading frame (ORF). It should be mentioned that plasmid pGM396 rearrangements have been found in a significant percentage of transformed cells, probably because of the toxicity of the hybrid RpsO-HA protein. On the contrary, plasmid rearrangements were never observed after transformation with plasmids pGM397 and pGM398. All fragments were obtained by polymerase chain reaction (PCR) on MG1655 (34) or λ genomic DNA with suitable oligonucleotides, sub-cloned by standard molecular biology techniques and checked by sequencing.

LD broth (35) was supplemented with chloramphenicol (30 μg/ml), ampicillin (100 μg/ml) and kanamycin (50 μg/ml) when needed.

Modulation of S1 expression in *E. coli* cultures

For S1 overexpression, strains harbouring both pQE31-S1 and pREP4 plasmids were grown at 37°C in LD broth in a reciprocating waterbath until $OD_{600} = 0.4$ was reached. The cultures were then split in two, 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to one of the subcultures and incubation was continued at 37°C. Samples were taken at different times for RNA extraction or crude extracts preparation. S1 depletion was achieved in the *araBp-rpsA* conditional expression mutant C-5699. The strain was grown in LD broth supplemented with 1% arabinose (permissive condition) at 37°C in a reciprocating waterbath up to $OD_{600} = 0.2$. The cells were then collected by centrifugation, washed with 1vol of LD and diluted 4-fold in LD with 0.4% glucose (non permissive-depletion condition). Incubation at 37°C was then

continued until the culture stopped growing (around OD_{600} = 0.4–0.5).

Northern blotting and primer extension

Procedures for RNA extraction, northern blot analysis, synthesis of radiolabelled riboprobes by *in vitro* transcription with T7 RNA polymerase and 5′-end labelling of oligonucleotides with T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP were previously described (36,37). As a loading control, urea-polyacrylamide gels were routinely stained with ethidium bromide before blotting and the intensity of the 5S rRNA band was evaluated with Quantity One (Bio-Rad) software. Coordinates of the CSPE riboprobe specific for *cspE* mRNA were 656576-656704. The oligonucleotide probes used in northern blotting experiments were: HA, 2135 (*cspE* ORF 3′-end), 2399 (*cspE* ORF 5′-end), 2313 (*rpsO* ORF), 2469 (*rpsO* leader) and 2521 (*rpsO* chromosomal allele). Autoradiographic images and densitometric analysis of northern blots were obtained by phosphorimaging using ImageQuant software (Molecular Dynamics). mRNA half-lives were estimated as described (6) by regression analysis of mRNA remaining (calculated as the densitometric signal at a given time) versus time after rifampicin addition.

Primer extension was performed on 10 µg of RNA extracted from different strains, as detailed in Figure 3B and Supplementary Figure S4 legends, with either the 5′-end ^{32}P-labelled 2174 (internal to *cspE* ORF) or HA oligonucleotides as previously described (38).

Analysis of protein and mRNA distribution in cell fractions

Escherichia coli cultures expressing S1 at different levels were grown as detailed above, whereas strains with autogenously regulated S1 were grown at 37°C in LD broth in a reciprocating waterbath until OD_{600} = 0.7–0.8 was reached. Preparation of crude extracts was performed as described by Charollais *et al*. (39) by freeze thawing in buffer A (10 mM Tris–HCl, pH 7.5, 60 mM KCl, 10 mM $MgCl_2$). The extract concentration was estimated by measuring the OD_{260}. Ribosomes and ribosomal subunits were prepared by centrifugation of the lysate at 100 000*g* for 2 h at 4°C. The supernatant (S100 fraction) was taken and the pellet was carefully washed and resuspended in one volume of buffer A. RNA was prepared by phenol–chloroform extraction from equal volumes of crude extract before ultracentrifugation (total), pellet (ribosomal fraction) and supernatant (S100). After ethanol precipitation of RNA, the samples were resuspended in identical volumes of RNase-free water.

To analyse the polysome profile, 14 OD_{260} units of the crude extracts were layered onto a 10–40% (w/v) sucrose gradient in 10 mM Tris–HCl, pH 7.5, 50 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM dithiothreitol (DTT) and centrifuged at 35 000 r.p.m. for 2.5 h at 4°C in a Beckman SW41 rotor. After centrifugation, 0.2 ml fractions were collected and their OD_{260} was plotted. The areas of the peaks corresponding to S100 (top of the gradients), ribosomal subunits, monosome and polysome fractions were estimated by weighting. Ten microlitres of selected fractions were then assayed by western blotting with antibodies specific for L4 ribosomal protein (kindly provided by C. Gualerzi) to confirm the correspondence of the peaks with ribosomal subunits, monosomes and polysomes, and with S1 specific antibodies (kindly provided by U. Bläsi). The immunoreactive bands were revealed by Immobilon (Millipore) reagents and quantified with the ImageQuant software. S1 densitometric values were normalized to the highest value obtained and plotted on the ribosomal profile chart. For *cspE* mRNA analysis, 20 OD_{260} of crude extracts were loaded on sucrose gradients and fractionated as described above. Each strain and condition tested was prepared in duplicate and ultracentrifuged together. Corresponding 0.3 ml fractions of duplicate gradients were pooled and the OD_{260} of the fractions measured. S1 was quantified in a subset of such samples by western blotting as described above. RNA was prepared by phenol–chloroform extraction of 0.3 ml of selected pooled fractions; after ethanol precipitation, the samples were resuspended in 0.15 vol of RNase-free water. Identical volumes of each RNA sample were analysed by northern blotting.

Electrophoretic mobility shift assay and *in vitro* degradation assays

The RNA probes used in electrophoretic mobility shift assay (EMSA) and *in vitro* degradation assays were synthesized by *in vitro* transcription of proper DNA fragments with T7 RNA polymerase following the protocol recommended by the enzyme manufacturer. The DNA templates were obtained by PCR on MG1655 genomic DNA with oligonucleotides 2299 (complementary to the 3′-part of *cspEt*) and oligonucleotides 2434 (T7 promoter + *cspE* 656473-656495) for *cspE*$^+$ or 2433 (T7 promoter + *cspE* 656453-656472) for Δ*L-cspE*. In order to obtain uniformly labelled probes, the reactions were carried out in the presence of [α-^{32}P] CTP. 5′-labelled RNA probes were prepared by dephosphorylation with alkaline phosphatase of unlabelled probes, followed by phosphorylation with T4 polynucleotide kinase and [γ-^{32}P] ATP. For gel retardation assays, 0.7 fmol of each probe was incubated for 20 min at 21°C in binding buffer [50 mM Tris–HCl pH 7.4, 50 mM NaCl, 0.5 mM DTT, 0.025% NP40 (Fluka), 10% glycerol] with increasing amounts of purified S1 protein in a final volume of 10 µl. The samples were run on 5% native polyacrylamide gel at 4°C. After run, the gel was dried and analysed by phosphorimaging. The signals were quantified using ImageQuant (Molecular Dynamics) software. For *in vitro* degradation experiments, 1.2 pmol of *in vitro* transcribed radiolabelled *cspE*$^+$ or Δ*L-cspE* RNA were incubated in 10 mM Tris–HCl pH 7.4, DTT 0.75 mM, Mg Acetate 4.5 mM, KCl 10 mM with 30 ng of RNA degradosome (prepared as described in ref. 40) at 26°C in a final volume of 15 µl. The experiment was performed in the absence of phosphate and NDPs, so as to prevent PNPase exonucleolytic and polymerization activities. Samples (3 µl) were removed at different time points and the reaction was stopped by adding 5 µl of RNA loading dye [2 mg/ml XC and BBF, 10 mM

ethylenediaminetetraacetic acid (EDTA) in formamide]. The samples were run on a 6% acrylamide denaturing gel. The gel was dried and analysed by phosphorimaging.

RESULTS

S1 ribosomal protein intracellular distribution is affected by S1 overexpression and depletion

Both S1 depletion and overexpression impair bacterial growth, presumably because altered S1 levels may affect translation (6,13). However, whereas translation inhibition is an expected outcome of S1 depletion, as S1 seems to be involved in translation initiation of most mRNAs (13), the effect of S1 overexpression on translation is far less obvious. To better understand the consequences of different S1 intracellular levels, we analysed the polysome profile and the intracellular distribution of S1 after modulating its expression from depletion to overexpression. Crude cell extracts were fractionated by ultracentrifugation on sucrose gradient and the polysome profile was assessed as described in 'Materials and Methods' section (Supplementary Table S2). Selected fractions were analysed by western blotting with anti-S1 antibodies and the signals quantified by densitometry. The results are shown in Figure 1A. In conditions of physiologically regulated S1 expression (S1 basal), S1 is present in the top (S100 S1), 30S, monosome and polysome fractions. S1 overexpression led to a drastic decrease of polysomes and ribosomal subunits and to accumulation of monosomes. A similar profile was observed in S1-depleted cells (S1 depletion), albeit in this case the polysome and ribosomal subunit peaks reduction was less severe. It should be reminded that in the latter condition S1 did not completely disappear even at late time points after shift to non-permissive conditions (6). Western blotting analysis of the fractions showed that the amount of S1 not associated to ribosomes (top fractions) clearly increased upon overexpression of the protein. On the contrary, upon S1 depletion, S1 was found only in 30S, monosome and residual polysomes fractions, whereas it was absent in the top fractions of the gradient (Figure 1A). We measured the amount of S1 relative to ribosomal protein L4 and to PNPase, taken as loading controls for 70S and free fractions, respectively, by western blotting of selected fractions (indicated by the arrows in Figure 1A). As can be seen in Figure 1B, variation of S1 expression strongly affected the non-ribosomal S1 pool; however, upon overexpression S1 slightly increased also in the monosomes whereas it sharply decreased in depleted cells (Figure 1B). It should be mentioned that the presence in the monosome fraction of PNPase, which was used in this experiment as a loading control, is presumably due to its association to the high molecular weight complex RNA degradosome, since in the *rne-131* mutant, which encodes a C-terminally truncated RNase E that does not assemble the degradosome (41,42), PNPase was found only in the top fractions (Supplementary Figure S1).

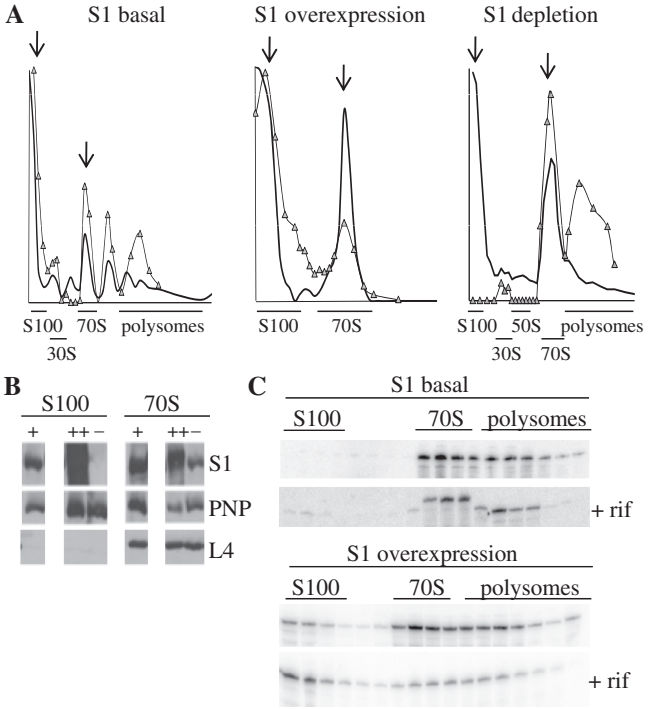


Figure 1. Ribosomal profile and intracellular distribution of S1 and *cspE* mRNA. Crude cell extracts were prepared as detailed in 'Materials and Methods' section from the following strains and conditions. S1 basal: C-1a exponential culture grown up to OD_{600} = 0.8; S1 overexpression: C-1a/pQE31S1/pREP4 was grown up to OD_{600} = 0.4 and incubated 60 min with 1 mM IPTG to induce *rpsA* transcription; S1 depletion: C-5699 (*araBp-rpsA*) grown up to OD_{600} = 0.2 in permissive conditions (LD +arabinose) was diluted 1 : 4 in non-permissive conditions (LD +glucose) to switch off *rpsA* transcription and further incubated for about 120 min. Cultures were grown at 37°C; before collecting the cells, the cultures were incubated 5 min at 37°C with chloramphenicol (final concentration, 0.1 mg/ml) to prevent polysome dissociation (39). Crude cell extracts (14 OD_{260}) were then fractionated by ultracentrifugation on 10–40% sucrose gradients. (**A**) Ribosomal profile. Thick continuous line: OD_{260} measured for each gradient fractions; grey triangles: S1 distribution. Ten microlitres of the indicated fractions were assayed by western blotting with S1-specific antibodies and the densitometric values (obtained as described in 'Materials and Methods' section) were normalized for the highest value obtained. (**B**) Distribution of S1, PNPase and ribosomal protein L4 in the S100 and 70S fractions. 0.02 OD_{260} of S100 and 70S fractions indicated by arrows in (A) were analysed by western blotting with specific antibodies (as indicated beside the panels). +, S1 basal; ++, S1 overexpression; −, S1 depletion. (**C**) Intracellular distribution of *cspE* mRNA. For extracts preparation, 100 ml of culture were taken before and 25 min after addition of 0.4 mg/ml rifampicin and 0.03 mg/ml nalidixic acid (+ rif). RNA was extracted from equal volumes of selected fractions; identical aliquots of RNA samples were loaded on a 6% denaturing polyacrylamide gel and analysed by northern blotting with the CSPE riboprobe. The altered electrophoretic mobility of *cspE* transcripts observed in the 70S samples (S1 basal, + rif) is probably imputable to the high concentration of ribosomal RNA in those fractions, since these transcripts migrated with the expected mobility upon sample dilution (data not shown).

S1 overexpression decreases the fraction of ribosome-associated *cspE* mRNA

To clarify whether mRNA stabilization upon S1 overexpression differentially affected ribosome-bound or unbound transcripts, we analysed the intracellular

distribution and stability of *cspE* mRNA relative to the ribosomes. We chose the *cspE* mRNA as a model because it is highly expressed and subject to S1-dependent stabilization (6). Crude extracts of S1 overexpressing and control (physiologically regulated *rpsA*) cultures immediately before and 25 min after rifampicin addition were fractionated on sucrose gradients. In the control gradients (Figure 1C, S1 basal), the *cspE* mRNA was found in monosome and polysome fractions; a tiny amount of mRNA (~0.2%; calculated by densitometry of the northern signals) was also detected upon longer exposure in the top fractions together with shorter RNAs, presumably *cspE* transcript degradation fragments (data not shown). After incubation with rifampicin, the polysomes disassembled and the 70S peak size increased (data not shown), in agreement with previously published data (43). Although most of the residual mRNA (on the whole, about one-tenth of the initial amount of RNA) was still located in the monosome and residual polysome fraction, the relative amount of the *cspE* transcript in the top versus ribosomal fractions increased almost 6-fold, as assessed by densitometric analysis of the signals (Figure 1C, S1 basal, +rif). In S1 over-expressing cells (Figure 1C, S1 overexpression), most mRNA appeared to be associated with monosomes and polysomes; however, the mRNA was readily detected in the top fractions before addition of rifampicin. Also in this case, after the incubation with the antibiotic, the relative amount of the *cspE* transcript at the top of the gradient significantly increased (from about 7% to 40%). It thus appears that the increased stability of *cspE* mRNA may be mainly attributed to the higher fraction of ribosome unbound transcripts.

S1 expression modulation differently affects leaderless and leadered transcript stability

It has been proposed that single-stranded regions in 5′-UTRs may constitute preferential S1 binding sites (22,23). To test whether the 5′-UTR was relevant for S1-dependent modulation of mRNA stability, we analysed in different conditions of S1 expression the transcript pattern of a *cspE* allele devoid of its leader region (Δ*L-cspE*) ectopically expressed from its natural *cspEp* promoter (plasmid pGM924, Figure 2A); as assessed by primer extension (Figure 3B), the Δ*L-cspE* primary transcript started with the gene start codon. In the northern blotting experiments shown in Figure 3A, we observed that when S1 expression was physiologically regulated, the leaderless transcript (Δ*L-cspE*) was much less stable than the leadered RNA ($cspE^+$) (half-life of <1 min for the Δ*L-cspE* versus about 4 min for the chromosomal $cspE^+$). Upon S1 overexpression, both transcripts were stabilized; in addition, several shorter RNAs were detected by the probe (Table 1 and Figure 3A). These short transcripts derived from degradation of the leaderless mRNA since they were not detectable in the wild-type strain lacking the plasmid (6), whereas they were produced in the Δ*cspE* chromosomal mutant carrying pGM924 (Figure 3A, right panel). Such short transcripts terminate downstream of the *cspE* stop codon, probably at the gene intrinsic terminator, and differ at their 5′-ends, as assessed by northern

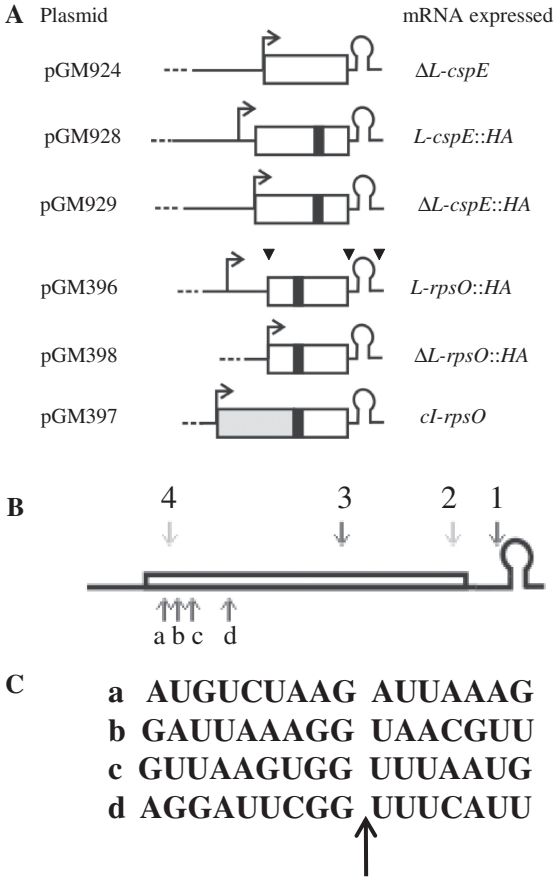


Figure 2. Map of plasmid-encoded *cspE* and *rpsO* alleles and of endonucleolytic cleavage sites on *cspE* mRNA. (**A**) Map of *cspE* and *rpsO* alleles cloned in pGM385 plasmid vector. Details about plasmid construction and coordinates of the cloned regions are reported in 'Materials and Methods' section. Transcription from *cspEp* and *rpsOp* promoters starts at nucleotide 657 473 and 3 309 808, respectively; coordinates of the palindromic region of *cspEt* and *rpsOt* transcription terminators are 656 744–656 768 and 3 309 420–3 309 394 (45,69). The promoters, 5′-UTRs and coding regions of the model genes are drawn to scale, whereas 3′-UTRs elements are reported on an arbitrary scale. Dotted line, vector sequence; bent arrow, promoters; hairpin, Rho-independent terminator; black box, HA epitope coding region. The *cspE* constructs carry the *cspEp* constitutive promoter. The grey box in pGM397 represents phage λ *cI* 5′ region. In pGM396 and pGM398, transcription of *rpsO*::HA alleles is driven by the *rpsOp* promoter; in pGM398, the 5′-UTR of *rpsO* was deleted and the transcript from *rpsOp* produced by the plasmid started with the A of the AUG of the gene (as assessed by primer extension; Supplementary Figure S4). In pGM397, the *rpsOp* and the 5′-end of the ORF, up to the HA insertion point, were replaced by P_{RM} and the first 63 codons of λ *cI* gene, which is naturally leaderless when transcribed from that promoter (70). The black triangles above pGM396 indicate the position of three RNase E cleavage sites mapped in *rpsO*, M2 (immediately upstream of the *rpsOt* terminator), M3 (at the beginning of *rpsO* coding sequence) and M sites (overlapping *rpsOt*) (69,71,72). (**B**) Map of endonucleolytic cleavage sites on *cspE* mRNA. Upward arrows: *in vivo* cleavage sites on Δ*L-cspE*; the 5′ ends of degradation products were mapped by primer extension (Figure 3B). Downward arrows: *in vitro* RNase E-cleavage sites detected on both leadered $cspE^+$ and Δ*L-cspE* (black arrows) or on either $cspE^+$ (site 4; data not shown) or Δ*L-cspE* (site 2) (grey arrows; see Figure 4B). (**C**) Nucleotide sequence associated with cut sites observed *in vivo*. The position of the listed sites in *cspE* mRNA is shown in (B) above. The arrow indicates the cleavage position.

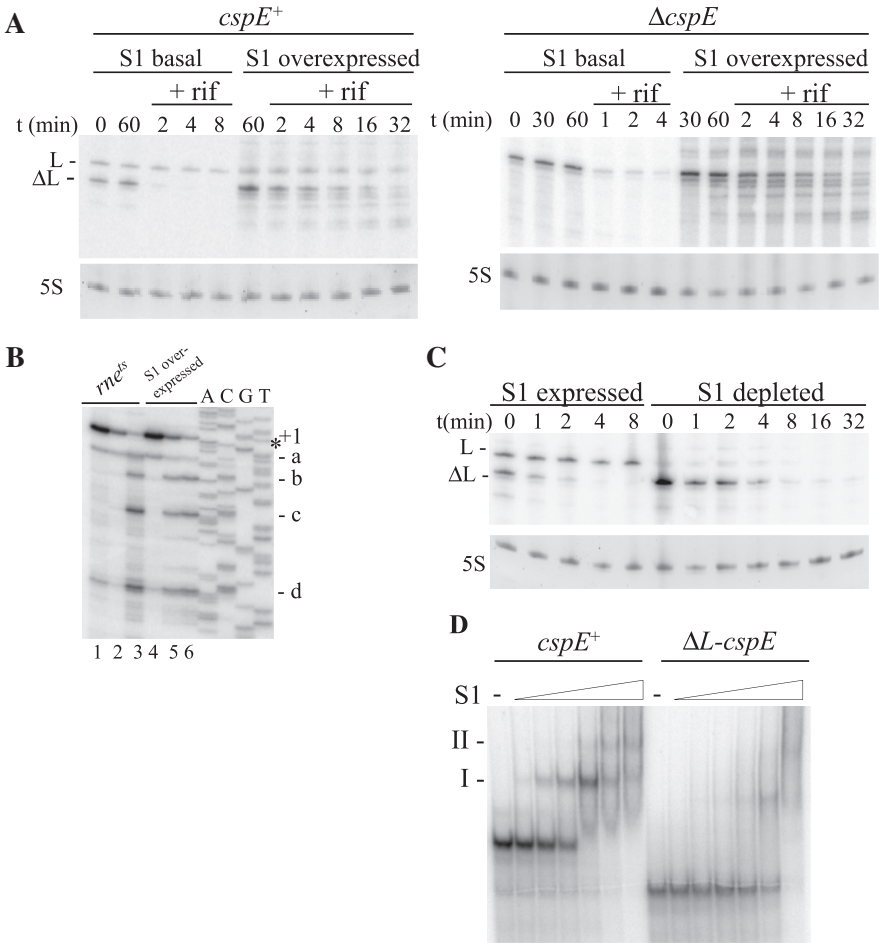


Figure 3. Analysis of leadered and leaderless *cspE* alleles transcripts. (**A**) Expression and stability of leadered and leaderless *cspE* mRNA upon S1 overexpression. Exponential cultures of C-1a/pREP4/pQE31-S1/pGM924 (*cspE*$^+$) or C-5874/pREP4/pQE31-S1/pGM924 (Δ*cspE*) ectopically expressing the leaderless *cspE* allele from pGM924 were grown up to OD_{600} = 0.4 (time = 0) and split in two subcultures. In one subculture (S1 over-expressed) S1 expression was induced by 1 mM IPTG addition and after 60 min (time = 60) the cultures were treated with rifampicin (0.4 mg/ml) and nalidixic acid (0.03 mg/ml). Aliquots for RNA extraction were sampled at times 0 and 60 (no antibiotics) and at different time points after addition of the antibiotics, as indicated (in min) on top of the lanes. Northern blotting was performed as described in 'Materials and Methods' section after 6 % denaturing polyacrylamide gel electrophoresis of 5 µg of RNA samples hybridized with radiolabelled CSPE riboprobe (upper panels). (Bottom panels) the gel was stained with ethidium bromide before transfer to check the amount of the loaded RNA samples. The gel portion with 5S rRNA signals is shown. L, leadered *cspE* chromosomal transcript; ΔL, leaderless *cspE* plasmid transcript. (**B**) Primer extension on leaderless *cspE* RNA degradation products. Selected RNA samples extracted from cultures of C-5868/pGM924 (grown as described in Figure 4A legend) and C-5874/pREP4/pQE31-S1/pGM924 (grown as described here above) were analysed by primer extension with oligonucleotide 2174, as described in 'Materials and Methods' section. The positions of different 5′-ends (a, +10; b, +18; c, +31 and d, +54) relatively to the first A of the primary transcript (+1) were defined by running the samples along with DNA sequencing reactions obtained with the same oligonucleotide and plasmid pGM924. The star beside the sequence indicates the A in the AUG initiation codon of *cspE* gene.1, C-5868/pGM924, 32°C, time 0; 2 and 3, 44°C, 0 and 4 min after rifampicin addition; 4, 5 and 6, C-5874/pREP4/pQE31-S1/pGM924 before (4) and 60 min after (5 and 6) S1 induction by IPTG. Samples 4 and 5 were taken before rifampicin addition (time 0), sample 6, 4 min after. (**C**) S1 depletion. Bacterial cultures of C-5699/pGM924 were grown up to OD_{600} = 0.4, diluted 1:4 in permissive (+ arabinose, S1 expressed) or non-permissive (+ glucose, S1 repressed) conditions and further incubated until OD_{600} = 0.4–0.5 was reached. RNA was extracted from rifampicin-nalidixic acid treated and untreated samples and northern blotted with radiolabelled 2135 oligonucleotide, as described above (upper panel). (Bottom panel) Ethidium bromide-stained 5S rRNA. L, leadered *cspE* chromosomal transcript; ΔL, leaderless *cspE* plasmid transcript. (**D**) EMSA with purified S1. Radiolabelled *cspE*$^+$ and Δ*L-cspE* RNAs were synthesized *in vitro* in the presence of [α^{32}P]-CTP. The probes (0.7 nM) were incubated 20 min at 21°C without (−) and with increasing amount of S1 (0.1, 0.5, 2.5, 5 and 25 nM). The samples were run on a 5% native polyacrylamide gel.

blotting with oligonucleotides specific for the 3′- and 5′-end (Supplementary Figure S2) and by primer extension analysis (Figure 3B). The position of the different 5′-ends detected and the nucleotide sequence of the cut sites are reported in Figure 2B and C, respectively.

Upon S1 depletion, the *cspE*$^+$ transcript became almost undetectable (6), whereas Δ*L-cspE* became more abundant and stable (Table 1 and Figure 3C). It should be noted that only the primary transcript but not the decay intermediates was stabilized in this condition.

We also assessed S1 ability to bind *in vitro* either the leadered or the leaderless *cspE* RNAs by EMSA. As shown in Figure 3D, His-tagged S1 formed two complexes with the *cspE*$^+$ transcript: complex I, which formed at S1 concentration as low as 0.1–0.5 nM, and complex II, which migrated more slowly than complex I and could

Table 1. mRNA expression and stability upon S1 depletion and overexpression

mRNA	S1 depletion[a]			S1 over-expression[a]		
	R.A.[b]	mRNA half-life (min)[c]		R.A.[b]	mRNA half-life (min)[c]	
		ARA	GLU		−IPTG	+IPTG
cspE^{+}[d]	0.2 ± 0.0	>4.0	2.5 ± 0.7	1.2 ± 0.1	7.2 ± 1.1	24.4 ± 4.7
Δ*L-cspE*[e]	2.9 ± 0.3	0.8 ± 0.3	2.0 ± 0.8	1.4 ± 0.5	0.6 ± 0.3	5.7 ± 0.5
rpsO^{+}[d]	0.2 ± 0.1	1.8 ± 0.2	1.6 ± 0.0	0.2 ± 0.1	1.8 ± 0.4	11.3 ± 2.1
Δ*L-rpsO::HA*[f]	3.1 ± 0.6	0.6 ± 0.0	1.0 ± 0.3	2.6 ± 0.2	1.2 ± 0.3	14.6 ± 5.8
cI-rpsO[g]	2.0 ± 0.4	1.0 ± 0.1	1.8 ± 0.1	3.2	1.0	9.4

[a]Cultures of C-5699 (S1 depletion) or C-1a/pQE31-S1/pREP4 (S1 overexpression) were grown and experiment performed as detailed in Figure 3 and Supplementary S5 legends and in 'Materials and Methods' section.
[b]Relative abundance, calculated as the ratio between mRNA amounts in cultures incubated with glucose and arabinose immediately before rifampicin addition (S1 depletion) or between induced and non-induced cultures 60 min after IPTG addiction (S1 overexpression).
[c]Calculated as detailed in 'Materials and Methods' section; the reported values represent average and standard deviation of at least two independent determinations in all cases, but the *cI-rpsO* mRNA in S1 overexpression, for which they are the results of a single determination.
[d]Leadered mRNAs. Cultures carrying either pGM924 (*cspE*) or pGM398 (*rpsO*) plasmid. mRNA expressed from chromosomally encoded alleles were considered for half-life calculation. For *rpsO*$^{+}$, only the mRNA terminated at *rpsOt* was considered. Albeit the stabilization factor for *rpsO*$^{+}$ mRNA by over-expressed S1 is in agreement with previous determinations performed in our laboratory, the half-life absolute values and especially the relative abundance (R.A.) reported here differ considerably from published data (6). This discrepancy is probably due to technical reasons, since data reported here refer to a single mRNA specie (the *rpsOp-rpsOt* mRNA) sharply resolved in polyacrylamide gel with very low background (see Supplementary Figure S5), whereas previously published data concerned the sum of two puffy agarose gel signals with high background (6).
[e,f,g]Leaderless alleles. Cells carrying pGM924 (e), pGM398 (f) or pGM397 (g) plasmid. Only signals corresponding to the primary transcripts (from *cspEp* to *cspEt*, e, *rpsOp-rpsOt*, f, and P_{RM}-*rpsOt*, g) expressed by the plasmids were considered for R.A. and half-life determination.

be detected at S1 concentrations ≥2.5 nM. Retardation of the leaderless transcript, on the contrary, occurred only at high (≥2.5 nM) S1 concentrations. This suggests that a high-affinity-binding site, responsible for complex I formation, may be present in the *cspE*$^{+}$ probe and missing in Δ*L-cspE*.

S1 inhibits RNase E-dependent decay *in vivo*

RNase E is thought to initiate the decay of several mRNAs by endonucleolytic cleavages at their 5′-end (44) and has been implicated in degradation of *cspE* mRNA (45,46). To test if Δ*L-cspE* mRNA decay products observed upon S1 overexpression could be generated by this enzyme, we analysed the transcription profile of the Δ*L-cspE* allele in an RNase E thermosensitive mutant. As shown in Figure 4A and summarized in Supplementary Table S3, in the *rne*^{ts} strain at 44°C the transcription profile of both the leadered (chromosomal) and the leaderless (plasmid-encoded) *cspE* alleles was strikingly similar to the pattern observed upon S1 overexpression, as both transcripts were stabilized and several shorter RNAs were detected after rifampicin addition, a condition that seems to exacerbate the RNase E-defective phenotype of *rne*^{ts} strains (47). These shorter decay products appeared to correspond to the Δ*L-cspE* decay intermediates that were stabilized in S1 over-expressing cells, since they: (i) were not detected in the absence of the plasmid pGM924; (ii) had identical 5′-ends, as assessed by primer extension, and exhibited the same electrophoretic mobility as those observed upon S1 overexpression; and (iii) hybridized with an oligonucleotide overlapping the *cspE* stop codon (Figure 3B and data not shown). Such degradation intermediates were still detected in the double *rne*^{ts} Δ*rng* mutant (lacking RNase G, a paralogue of RNase E with the same propensity to cleave within AU-rich single-stranded segments) (48), and in the double *rne*^{ts} Δ*elaC* mutant, lacking a functional RNase Z previously implicated in *cspE* mRNA decay (46) (Supplementary Figure S3), thus ruling out that these two nucleases are implicated in generation of such decay intermediates. The alignment of cleavage sites shows that the cut invariably occurs after a G residue embedded in a U-rich sequence. To our knowledge, none of the known *E. coli* endoribonucleases exhibits such cleavage specificity. Narrow substrate specificity is typical of mRNA interferases, which are components of toxin–antitoxin systems (49). S1 is known to stimulate the activity of bacteriophage T4 RegB endoribonuclease (50,51) which has structural similarities with two *E. coli* interferases, YoeB and RelE (52).

It thus appears that the leaderless Δ*L-cspE* transcript may be degraded by two pathways: an RNase E-dependent pathway that is inhibited by S1 overexpression and an alternative processing by an unidentified endonuclease, which is not prevented by S1. It will be interesting to identify such endonuclease and to analyse whether S1 may play a positive role in this degradation pathway.

To identify *in vitro* RNase E cleavage sites in both leadered *cspE* and Δ*L-cspE*, we digested these RNAs with purified RNA degradosome in conditions that prevented PNPase enzymatic activities. We performed this experiment with either 5′-radiolabelled (Figure 4B, left) or (Figure 4B, right) uniformly labelled probes. Monophosphorylated *cspE* and Δ*L-cspE* were degraded about 2-fold faster than ppp-mRNAs (assessed by plotting the amount of full-length substrate remaining at each time point; data not shown), suggesting that the degradation rate of our probes was only marginally affected by their phosphorylation state (53–55). Two main decay fragments

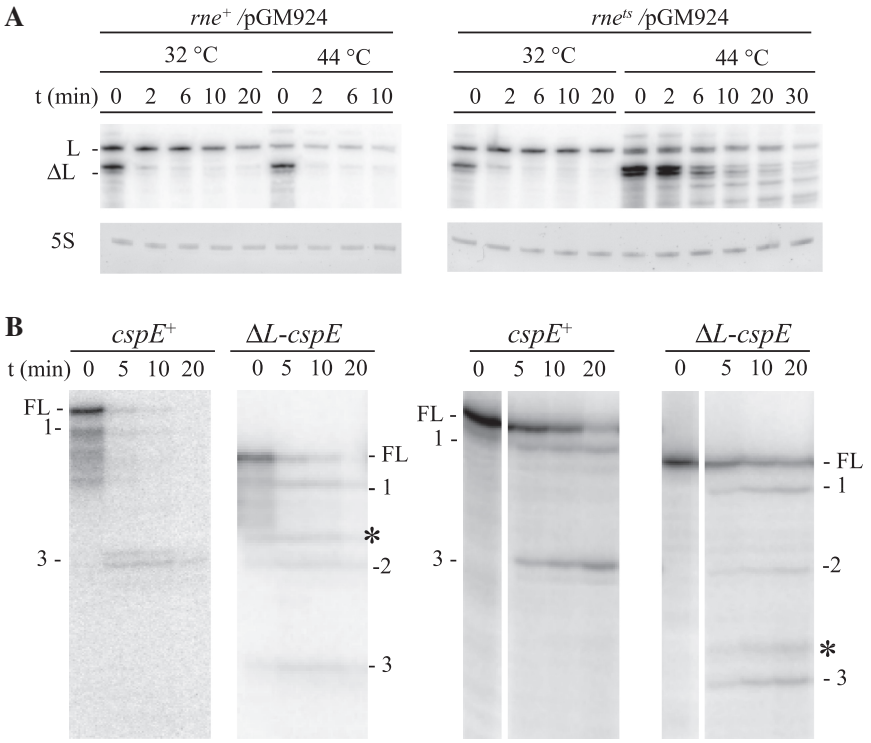


Figure 4. Analysis of RNase E role in $cspE^+$ and ΔL-*cspE* degradation. (**A**) *In vivo* analysis. Cultures of rne^+ (C-5869) and rne^{ts} (C-5868) strains carrying pGM924 were grown to mid-log phase at permissive temperature (32°C) and shifted at non-permissive temperature (44°C). Rifampicin was added immediately before (32°C samples) and 30 min after temperature shift (44°C) and RNA was extracted at the time points indicated on top of the lanes. Five micrograms of RNA samples were analysed by northern blotting with radiolabelled 2135 oligonucleotide. L, leadered *cspE* chromosomal transcript; ΔL, leaderless *cspE* plasmid transcript (upper panels). (Bottom panels) ethidium bromide-stained 5S rRNA. (**B**) *In vitro* degradation assay. $cspE^+$ and ΔL-*cspE* RNAs (35 nM) 5′-end ^{32}P-labelled (left) or uniformly radiolabelled with [α^{32}P]-CTP (right) were incubated at 26°C for the time indicated (in min) above the lanes with 60 ng of purified RNA degradosome and fractionated by 6% PAGE. The size of the main RNA species was estimated by running the samples along with a sequence ladder (data not shown). The corresponding leadered and leaderless RNAs are denoted by the same figure; their respective size (in nucleotides) and 3′-ends (coordinates from NCBI Accession Number U00096.2.) are as follows: 1: 270/229, 656742; 2: 169 (leaderless only), 656 683; 3: 165/127, 656 638. The stars indicate signals that were not reproducibly detected in other experiments. Shorter decay fragments migrating at the bottom of the gel (not shown in the figure), probably corresponding to the probes 3′-end fragments, were present in the right panel.

were visible. The 3′-end of these fragments were mapped (by comparison with a sequencing ladder; data not shown) immediately upstream of the intrinsic terminator *cspEt* and internally to the ORF, respectively (Figures 2B and 4B, signals 1 and 3). An additional 3′-end internal to the ORF was generated only with ΔL-*cspE* RNA (Figures 2B and 4B, signal 2). Other signals occasionally observed (Figure 4B, stars) may represent unstable degradation intermediates. Among them, an additional cut occurring 50/51-nt downstream of the 5′-end was detected with the *cspE* probe (Figure 2B, signal 4; data not shown). Signals corresponding to RNA molecules processed at upstream sites were never observed. Therefore, *in vitro* RNase E cleavage sites and *in vivo* cuts detected upon S1 overexpression or RNase E thermal inactivation map in different regions of the leaderless *cspE* mRNA (Figure 2B). Addition of purified His-tagged S1 to the *in vitro* degradation assay did not consistently and reproducibly inhibit RNase E cleavage (data not shown). This could be due to technical limitations of our assay or it may suggest that other factors are needed to fully reconstruct the process *in vitro*.

S1 overexpression and depletion effects on mRNA are not specific for *cspE*

The data reported above show that ΔL-*cspE* may be degraded by both RNase E-dependent and independent pathways and that S1 overexpression may inhibit RNase E-dependent degradation. To clarify whether RNA protection by S1 was specific for this transcript or could be a more general phenomenon, we analysed *in vivo* other leaderless artificial mRNAs. The *rpsO* gene was used as a backbone of our constructs because it is sensitive to S1 stabilization (6) and its RNase E-dependent degradation pathway has been extensively studied (56). Two constructs were analysed, a leaderless *rpsO* gene (ΔL-*rpsO*::HA) and a λ *cI-rpsO* fusion (carried by plasmid pGM398 and pGM397, respectively; Figure 2A). In both constructs, the region encoding S15 binding site for 16S rRNA (33) was replaced with an in frame HA epitope (Figure 2A). This allowed discriminating transcripts deriving from chromosomal or plasmid alleles with specific probes in northern hybridization.

RNase E was expected to be involved in the decay of the ΔL-*rpsO*::*HA* transcripts, as all known RNase E cleavage

sites mapping in *rpsO* (Figure 2A) are conserved in this construct. To assess if this was indeed the case, we analysed the transcription profile of Δ*L-rpsO*::*HA* construct in the RNase E defective strain. We observed that the transcript covering the region between *rpsOp* and *rpsOt* (Supplementary Figure S5A, P-t, and Supplementary Table S3) and longer RNAs deriving from transcriptional read-through of the *rpsOt* (Supplementary Figure S5A, stars) were clearly stabilized in the *rne*^{ts} strain at 44°C. Thus, Δ*L-rpsO*::*HA* RNA appears to be degraded via an RNase E-dependent pathway. It is very likely that the same holds true also for *cI-rpsO*, as two known RNase E-dependent cleavage sites located respectively immediately upstream and overlapping the *rpsOt* terminator are present in pGM397 construct (Figure 2A).

We then analysed the effect of S1 expression modulation on *rpsO* transcription profile. Leadered *rpsO*$^{+}$ mRNA expressed by the chromosomal allele was stabilized six to seven times by over-expressed S1 in these strains (Supplementary Figure S5B and Table 1). Leaderless transcripts deriving from Δ*L-rpsO*::*HA* and *cI-rpsO* constructs were quite unstable, with half-lives around 1 min, and were strongly stabilized by S1 overexpression (Supplementary Figure S5 and Table 1); in the latter condition, the amount and stability of RNAs expressed by the two constructs that encompass the *rpsOt* intrinsic terminator were also clearly increased (Supplementary Figure S5B, upper panel, stars). Thus, as for Δ*L-cspE* transcripts, overexpression of S1 stabilizes these otherwise very unstable leaderless RNAs by inhibiting their RNase E-dependent decay.

As observed for *cspE*, upon S1 depletion the two leaderless transcripts became 2–3-fold more abundant and slightly more stable, whereas the amount of both the leadered *L-rpsO*::*HA* expressed by pGM396 and *rpsO*$^{+}$ chromosomal transcript decreased to 50–60% and 20%, respectively, of the quantity present before depletion. However, the half-life of the remaining *rpsO*$^{+}$ mRNA did not significantly change irrespective of S1 expression level (Supplementary Figure S5C and Table 1).

Leaderless transcripts co-localize with ribosomes irrespective of translation

The above-mentioned results show that S1 expression modulation, and depletion in particular, differentially affects leadered and leaderless RNA expression and stability. It is possible that differences in ribosome association and/or translation efficiency may contribute to this difference. We thus assayed the association with ribosomes of leadered and leaderless RNAs expressed by the above constructs (Figure 2) in different conditions of S1 expression. To do so, RNA from crude extracts, ribosomal fractions and S100 fractions prepared from different cultures (listed in Figure 5 legend), as described in 'Materials and Methods' section, were analysed by northern blotting (Figure 5). Leadered transcripts were found only in the ribosomal fraction at physiological levels of S1 expression; after S1 overexpression, they were present both in the ribosomal and in the S100 fractions. Leaderless mRNAs were found mainly in the ribosomal fraction before S1 induction and only in the S100 upon S1 overexpression, with the exception of *cI-rpsO* transcript that associated to the ribosomal fraction irrespective of S1 expression levels. This could be explained by the presence in *cI* mRNA of an out of frame AUG (nucleotides 68–70) preceded by a properly positioned SD sequence (57).

Interestingly, S1 depletion did not affect ribosome association of the leaderless transcripts (Figure 5). Moreover, we assayed by western blotting with anti-HA antibodies the translation efficiency of the leadered and leaderless set of constructs tagged with the HA epitope shown in Figure 2. We found that both *L-cspE*::*HA* and *L-rpsO*::*HA* leadered transcripts were translated; the expression of both proteins was sensitive to S1 levels, as it sharply decreased in S1 depletion and, to a lesser extent (about 60% of pre-induction level for *L-rpsO*::*HA* and 50% for *L-cspE*::*HA*), upon S1 overexpression (Figure 6). The leaderless Δ*L-cspE*::*HA* and Δ*L-rpsO*::*HA* transcripts, albeit co-localizing with ribosomes (Figure 5), were detectably translated in none of the different conditions of S1 expression. On the contrary, the amount of CI-RpsO hybrid protein (relative to L4 ribosomal protein amount) did not change in any condition tested, irrespective of S1 levels (Figure 6).

DISCUSSION

S1 protein can act as a negative modulator of translation and RNase E-dependent mRNA decay

In this work, we have investigated S1 role in translation and mRNA stability control by altering S1 expression level. Our data show that S1 overexpression causes polysome disappearance and translation inhibition. Moreover, a sharp increase in the amount of S100 S1 (ribosome-unbound) and cellular re-distribution of mRNA from ribosomal to S100 fractions are observed (see Figures 1C and 5). This suggests that *in vivo* the ribosome-unbound S1 can negatively affect the association between ribosome and mRNA, as already demonstrated *in vitro* for different mRNAs (58). Indeed, if the first step in initiation complex formation requires the interaction between 30S-associated S1 and the mRNA leader region (17,18), S1 binding to high-affinity site(s) in the mRNA leader regions may hamper mRNA binding to S1-containing 30S subunits (22,23) (see Figure 3D). Translation repression by S1 expressed at physiological level has been documented for a couple of *E. coli* genes. In fact, S1 acts as a repressor of its own gene translation by preventing association of its own mRNA to 30S (58,59). *rpsA* mRNA lacks a canonical SD sequence; for this messenger, 30S recruitment could be strictly mediated by 30S-bound S1 and thus efficiently counteracted by free (ribosome-unbound) S1. S1 has been also claimed to repress translation of the dicistronic *rpsB* and *tsf* (encoding elongation factor Ts) mRNA in cooperation with another ribosomal protein, S2 (60). Overexpression appears thus to intensify and extend an otherwise specific activity of S1 as a negative modulator of translation initiation.

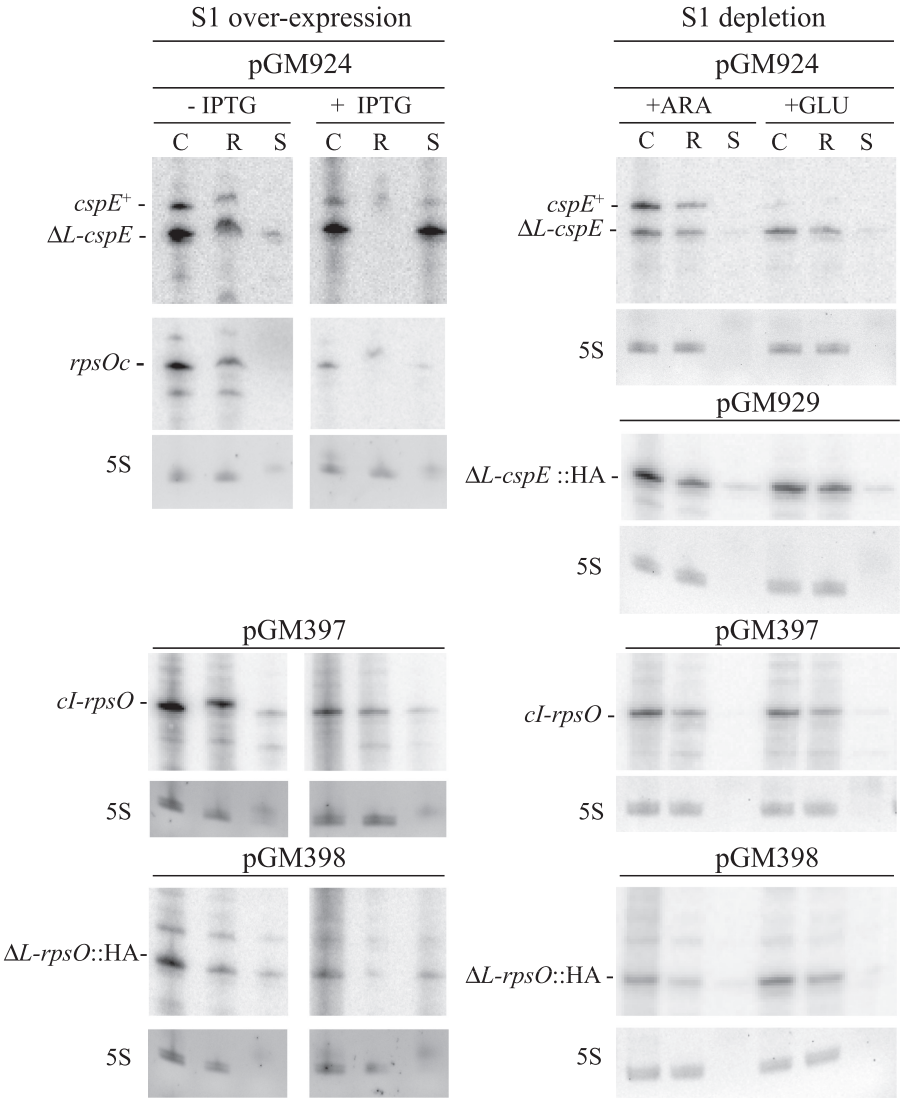


Figure 5. Intracellular distribution of leadered and leaderless mRNAs upon modulation of S1 expression. Northern blotting of RNA samples extracted from crude extracts (C) and from ribosomal (R) and S100 (S) fractions. Cultures of C-1a/pREP4/pQE31-S1 (S1 overexpression) or C-5699 (S1 depletion) strains carrying the additional plasmids listed above the panels were grown as detailed in 'Materials and Methods' section. Identical volumes of the RNA samples were analysed by 6% denaturing polyacrylamide gel electrophoresis, northern blotted and hybridized with the following oligonucleotides. pGM924, oligo 2135 (*cspE* panels) or 2469 (specific for *rpsO* mRNA expressed by the chromosomal gene; *rpsOc* panel); pGM397, pGM398 (S1 depletion) and pGM929, oligo HA; pGM398 (S1 overexpression), oligo 2313 (specific for *rpsO*). Left part, −IPTG, S1 basal; + IPTG, S1 induced. Right part, + ARA, S1 expressed, + GLU, S1 depleted.

Besides preventing the initial association between the mRNA and the ribosome, S1 overexpression may affect translation through a different and still puzzling mechanism. In fact, monosomes accumulation in S1 overexpressing cells suggests that either ribosomal subunits do not dissociate after releasing the mRNA or the 70S–mRNA complex remains associated. The presence of residual leadered mRNA in the ribosomal fractions after S1 induction (see Figure 1C) supports the latter hypothesis. At the moment, we have no experimental hint on a possible mechanism for mRNA trapping on ribosome. One can speculate that S1 dissociates from the ribosome at some point during translation and that overexpression may alter its cycling, preventing mRNA release. Evidences based on S1 stoichiometry determination in polysomes and on its role in translation elongation argue against S1 dissociation during elongation (9,13,61). On the other hand, leaderless mRNAs can be translated *in vitro* by S1-depleted 70S (14). Moreover, to our knowledge, S1 fate during translation termination and ribosome release has never been addressed. The existence of a ribosome-unbound S1 pool, whose magnitude depends on S1 expression level, and the observation that upon S1 overexpression 70S–mRNA–S1 complexes may be stabilized, support S1 recycling at physiological S1 concentration.

In spite of translation inhibition, different transcripts are stabilized by S1 overexpression. This is not due to a general impairment of RNA degradation, since Δ*L-cspE* degradation from the 5′-end by an unknown endonuclease

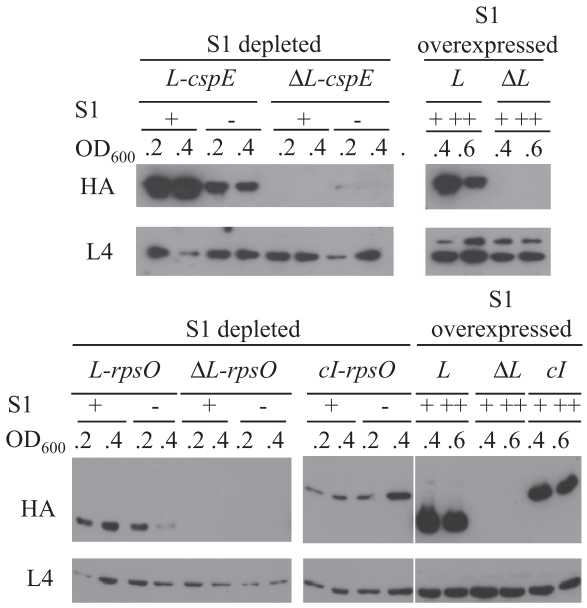


Figure 6. Translation of leadered and leaderless mRNAs upon modulation of S1 expression. Proteins extracted from cultures of C-1a/pREP4/pQE31-S1 (S1 depleted; +, S1 basal; ++, S1 induced) or C-5699 (S1 depleted; +, S1 expressed; −, S1 repressed) strains carrying *cspE*::*HA* (upper panels: *L-cspE*, pGM928; Δ*L-cspE*, pGM929) or *rpsO*::*HA* (lower panels: *L-rpsO*, pGM396; Δ*L-rpsO*, pGM398; *cI-rpsO*, pGM397) plasmids were prepared as detailed in 'Materials and Methods' section. Proteins were separated by 15% denaturing polyacrylamide gel electrophoresis and immunodecorated with antibodies specific for the HA epitope or, as a loading control, for L4 ribosomal protein. For quantitative evaluation of CspE::HA and RpsO::HA expression, both HA- and L4-specific antibodies signals were quantified with ImageQuant (Molecular Dynamics); each HA-specific signal volume was then normalized by the volume of the corresponding L4-specific signal.

is not prevented by S1. Conversely, S1 overexpression inhibits RNase E-dependent decay of our reporter transcripts. Another ribosomal protein, L4, has been involved in mRNA stability control, as it was shown to physically interact with RNase E and negatively regulate its endoribonucleolytic function (62). We cannot rule out that S1 may exert its inhibitory function by physically interacting with RNase E; however, we have not found RNase E among proteins co-purifying with His-tagged S1 (Briani,F., unpublished data). In addition, S1 has been shown to preferentially bind mRNA at A/U- rich single-stranded regions that also constitute potential RNase E cleavage sites (1,22–24). This suggests that S1 may inhibit RNase E decay by directly shielding cleavage site(s).

Our *in vitro* data (see Figure 3D) and data by other groups (22,23,63) suggest that S1 not bound to ribosomes may preferentially interact with sites located in the mRNA leader regions. This interaction may have different consequences on RNA stability depending on the specific decay pathway run by an RNA specie [(64); current models of RNase E degradation modes have been recently reviewed by refs. 44 and 65]. For RNAs degraded predominantly through a wave of cleavages by RNase E moving 5′ to 3′, S1 binding to the 5′-UTR could hinder the first RNase E cut, which is the rate-limiting step in this degradation pathway. This would uncouple translation and decay by protecting the translationally silenced mRNA from degradation. It would be interesting to investigate whether this is indeed the case for *rpsA* and *rpsB-tsf* mRNAs. More in general, in conditions of impaired translation, such as amino acid starvation or cold shock, ribosome-unbound S1 may protect part of cellular transcripts from decay. Best candidates for S1 protection would be transcripts with high-affinity sites for S1 in their leader regions and degraded 5′ to 3′ by RNase E, because they could be sensitive to a relatively modest free S1 increase. Conversely, free S1 binding to the 5′-UTR may result in very fast (possibly co-transcriptional) degradation of mRNAs attacked by RNase E at internal sites, because these sites will be no longer hindered by translating ribosomes. However, S1 binding at ectopic sites (i.e. downstream of the leader region) would allow some molecules to escape degradation, thus stabilizing a share of the mRNA. It has been proposed that RNase E may degrade *rpsO* mRNA predominantly through the 'internal entry' mode (65). Our *in vitro* data suggest that also *cspE* RNA may enter such degradation pathway, since the phosphorylation state of the 5′-end of our probes or even the presence of a leader region seem to marginally affect the *in vitro* degradation efficiency by RNase E. This may explain why, despite the strong half-life increase, the abundance of tested mRNAs does not correspondingly increase or even decreases in S1 over-expressing cells (Table 1) (6).

Ribosomal particles lacking S1 may be present in normally growing cells

Our artificial leaderless Δ*L-cspE* and Δ*L-rpsO*::*HA* RNAs co-fractionate with ribosomes at basal S1 expression, whereas they are present only in S100 fraction in S1 over-expressing cells. Their re-distribution upon S1 overexpression is more drastic than for leadered RNAs, which are, in part, still retained in ribosomal fractions, and may occur through a different mechanism.

Leaderless transcripts enjoy a peculiar translation initiation pathway based on the initial interaction of the terminal AUG with 70S particles (14,20,21). S1 has a documented destabilizing effect on this interaction (14,66). The presence of leaderless mRNAs in the ribosomal fraction at S1 physiological levels suggests that ribosomal particles without S1 may be present in the cells. The slight increase in S1 stoichiometry observed in monosome fractions upon overexpression also favours the idea that S1 could be normally present in slightly sub-stoichiometric amount in 70S particles, as previously suggested (9), and may reach stoichiometry when over-expressed. A conclusive experimental demonstration of this hypothesis would be very challenging, since S1 dissociation from ribosomal particles may occur after cell lysis (10–12). However, the idea that leaderless mRNAs may interact with 70S particles lacking S1 is strengthened by S1 depletion data. In this condition, 70S particles lacking S1 accumulate, probably because of their unusual stability (20,67), and also leaderless mRNAs, which are detectable only in ribosomal fraction, become more abundant, suggesting that in

this condition the share of RNA molecules stabilized by ribosome binding may increase.

In spite of their association with ribosomes, artificial leaderless mRNAs are not detectably translated in any assayed condition, suggesting that these transcripts may establish an unproductive interaction with 70S ribosome. On the contrary, the chimeric *cI-rpsO* mRNA, which contains the 5′-end of the naturally leaderless λ *c*I RNA, is translated both at physiological S1 expression and upon depletion (no conclusion can be drawn about *cI-rpsO* mRNA translation upon overexpression as the 60-min time of S1 induction is likely insufficient to get rid of the protein synthesized before the induction). In agreement with our data, it has been recently reported that a *cI-lacZ* fusion is translated 60-fold more efficiently than a leaderless *lacZ* reporter construct (68). The molecular bases of λ *cI* mRNA properties still remain elusive. However, it has been pointed out that to initiate translation, the terminal AUG of leaderless RNAs should be located at the ribosomal P-site. This would require the RNA to worm its way through the channel between the subunits of the 70S ribosome. The propensity of different transcripts 5′-end to fold into (stable) secondary structures may modulate their ability to enter into the channel and establish a fruitful interaction with the ribosome (14).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We thank U. Bläsi and C.O. Gualerzi for providing antibodies, M.V. Sukhodolets for providing plasmids, and the ''National BioResource Project (NIG, Japan): *E. coli*'' for providing bacterial trains. We thank Isabella Moll for critical reading of the manuscript.

FUNDING

Ministero dell'Università e della Ricerca and Università degli Studi di Milano (PRIN 2007); and Università degli Studi di Milano (FIRST2007 and PUR2008). Funding for open access charge: Università degli Studi di Milano.

Conflict of interest statement. None declared.

REFERENCES

1. Kaberdin,V.R. and Bläsi,U. (2006) Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol. Rev.*, **30**, 967–979.
2. Deana,A. and Belasco,J.G. (2005) Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.*, **19**, 2526–2533.
3. Draper,D.E., Pratt,C.W. and von Hippel,P.H. (1977) *Escherichia coli* ribosomal protein S1 has two polynucleotide binding sites. *Proc. Natl Acad. Sci. USA*, **74**, 4786–4790.
4. Kalapos,M.P., Paulus,H. and Sarkar,N. (1997) Identification of ribosomal protein S1 as a poly(A) binding protein in *Escherichia coli*. *Biochimie*, **79**, 493–502.
5. Feng,Y., Huang,H., Liao,J. and Cohen,S.N. (2001) *Escherichia coli* poly(A)-binding proteins that interact with components of degradosomes or impede RNA decay mediated by polynucleotide phosphorylase and RNase E. *J. Biol. Chem.*, **276**, 31651–31656.
6. Briani,F., Curti,S., Rossi,F., Carzaniga,T., Mauri,P. and Dehò,G. (2008) Polynucleotide phosphorylase hinders mRNA degradation upon ribosomal protein S1 overexpression in *Escherichia coli*. *RNA.*, **14**, 2417–2429.
7. Held,W.A., Mizushima,S. and Nomura,M. (1973) Reconstitution of *Escherichia coli* 30S ribosomal subunits from purified molecular components. *J. Biol. Chem.*, **248**, 5720–5730.
8. Subramanian,A.R. (1983) Structure and functions of ribosomal protein S1. *Prog. Nucleic Acid Res. Mol. Biol.*, **28**, 101–142.
9. van Knippenberg,P.H., Hooykaas,P.J. and van Duin,J. (1974) The stoichiometry of *E. coli* 30S ribosomal protein S1 on *in vivo* and *in vitro* polyribosomes. *FEBS Lett.*, **41**, 323–326.
10. Szer,W., Hermoso,J.M. and Leffler,S. (1975) Ribosomal protein S1 and polypeptide chain initiation in bacteria. *Proc. Natl Acad. Sci. USA*, **72**, 2325–2329.
11. Robertson,W.R., Dowsett,S.J. and Hardy,S.J. (1977) Exchange of ribosomal proteins among the ribosomes of *Escherichia coli*. *Mol. Gen. Genet.*, **157**, 205–214.
12. Subramanian,A.R. and van Duin,J. (1977) Exchange of individual ribosomal proteins between ribosomes as studied by heavy isotope-transfer experiments. *Mol. Gen. Genet.*, **158**, 1–9.
13. Sørensen,M.A., Fricke,J. and Pedersen,S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli in vivo*. *J. Mol. Biol.*, **280**, 561–569.
14. Moll,I., Grill,S., Gualerzi,C.O. and Bläsi,U. (2002) Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol. Microbiol.*, **43**, 239–246.
15. Kaberdina,A.C., Szaflarski,W., Nierhaus,K.H. and Moll,I. (2009) An unexpected type of ribosomes induced by kasugamycin: a look into ancestral times of protein synthesis? *Mol. Cell*, **33**, 227–236.
16. Tedin,K., Resch,A. and Bläsi,U. (1997) Requirements for ribosomal protein S1 for translation initiation of mRNAs with and without a 5′ leader sequence. *Mol. Microbiol.*, **25**, 189–199.
17. Hauryliuk,V. and Ehrenberg,M. (2006) Two-step selection of mRNAs in initiation of protein synthesis. *Mol. Cell*, **22**, 155–156.
18. Studer,S.M. and Joseph,S. (2006) Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol. Cell*, **22**, 105–115.
19. Farwell,M.A., Roberts,M.W. and Rabinowitz,J.C. (1992) The effect of ribosomal protein S1 from *Escherichia coli* and *Micrococcus luteus* on protein synthesis *in vitro* by *E. coli* and *Bacillus subtilis*. *Mol. Microbiol.*, **6**, 3375–3383.
20. Moll,I., Hirokawa,G., Kiel,M.C., Kaji,A. and Bläsi,U. (2004) Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. *Nucleic Acids Res.*, **32**, 3354–3363.
21. Udagawa,T., Shimizu,Y. and Ueda,T. (2004) Evidence for the translation initiation of leaderless mRNAs by the intact 70S ribosome without its dissociation into subunits in eubacteria. *J. Biol. Chem.*, **279**, 8539–8546.
22. Sengupta,J., Agrawal,R.K. and Frank,J. (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc. Natl Acad. Sci. USA*, **98**, 11991–11996.
23. Boni,I.V., Isaeva,D.M., Musychenko,M.L. and Tzareva,N.V. (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res.*, **19**, 155–162.
24. Komarova,A.V., Tchufistova,L.S., Dreyfus,M. and Boni,I.V. (2005) AU-rich sequences within 5′ untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *J. Bacteriol.*, **187**, 1344–1349.
25. Sasaki,I. and Bertani,G. (1965) Growth abnormalities in Hfr derivatives of *Escherichia coli* strain C. *J. Gen. Microbiol.*, **40**, 365–376.
26. Carzaniga,T., Briani,F., Zangrossi,S., Merlino,G., Marchi,P. and Dehò,G. (2009) Autogenous regulation of *Escherichia coli* polynucleotide phosphorylase expression revisited. *J. Bacteriol.*, **191**, 1738–1748.
27. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA*, **97**, 6640–6645.

28. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.
29. Sukhodolets,M.V. and Garges,S. (2003) Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry*, **42**, 8022–8034.
30. Zangrossi,S., Briani,F., Ghisotti,D., Regonesi,M.E., Tortora,P. and Dehò,G. (2000) Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in *Escherichia coli*. *Mol. Microbiol.*, **36**, 1470–1480.
31. Lessl,M., Balzer,D., Lurz,R., Waters,V.L., Guiney,D.G. and Lanka,E. (1992) Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region in trans. *J. Bacteriol.*, **174**, 2493–2500.
32. Briani,F., Zangrossi,S., Ghisotti,D. and Dehò,G. (1996) A Rho-dependent transcription termination site regulated by bacteriophage P4 RNA immunity factor. *Virology*, **223**, 57–67.
33. Ehresmann,C., Ehresmann,B., Ennifar,E., Dumas,P., Garber,M., Mathy,N., Nikulin,A., Portier,C., Patel,D. and Serganov,A. (2004) Molecular mimicry in translational regulation: the case of ribosomal protein S15. *RNA Biol.*, **1**, 66–73.
34. Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F. *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1462.
35. Ghisotti,D., Chiaramonte,R., Forti,F., Zangrossi,S., Sironi,G. and Dehò,G. (1992) Genetic analysis of the immunity region of phage-plasmid P4. *Mol. Microbiol.*, **6**, 3405–3413.
36. Dehò,G., Zangrossi,S., Sabbattini,P., Sironi,G. and Ghisotti,D. (1992) Bacteriophage P4 immunity controlled by small RNAs *via* transcription termination. *Mol. Microbiol.*, **6**, 3415–3425.
37. Briani,F., Del Favero,M., Capizzuto,R., Consonni,C., Zangrossi,S., Greco,C., De Gioia,L., Tortora,P. and Dehò,G. (2007) Genetic analysis of polynucleotide phosphorylase structure and functions. *Biochimie*, **89**, 145–157.
38. Forti,F., Sabbattini,P., Sironi,G., Zangrossi,S., Dehò,G. and Ghisotti,D. (1995) Immunity determinant of phage-plasmid P4 is a short processed RNA. *J. Mol. Biol.*, **249**, 869–878.
39. Charollais,J., Pflieger,D., Vinh,J., Dreyfus,M. and Iost,I. (2003) The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in *Escherichia coli*. *Mol. Microbiol.*, **48**, 1253–1265.
40. Mauri,P. and Dehò,G. (2008) A proteomic approach to the analysis of RNA degradosome composition in *Escherichia coli*. *Methods Enzymol.*, **447**, 99–117.
41. Lopez,P.J., Marchand,I., Joyce,S.A. and Dreyfus,M. (1999) The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing *in vivo*. *Mol. Microbiol.*, **33**, 188–199.
42. Leroy,A., Vanzo,N.F., Sousa,S., Dreyfus,M. and Carpousis,A.J. (2002) Function in *Escherichia coli* of the non-catalytic part of RNase E: role in the degradation of ribosome-free mRNA. *Mol. Microbiol.*, **45**, 1231–1243.
43. Blundell,M.R. and Wild,D.G. (1971) Altered ribosomes after inhibition of *Escherichia coli* by rifampicin. *Biochem. J.*, **121**, 391–398.
44. Carpousis,A.J., Luisi,B.F. and McDowall,K.J. (2009) Endonucleolytic initiation of mRNA decay in *Escherichia coli*. *Prog. Mol. Biol. Transl. Sci.*, **85**, 91–135.
45. Uppal,S., Akkipeddi,V.S. and Jawali,N. (2008) Post-transcriptional regulation of *cspE* in *Escherichia coli*: involvement of the short 5′-untranslated region. *FEMS Microbiol. Lett.*, **279**, 83–91.
46. Perwez,T. and Kushner,S.R. (2006) RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Mol. Microbiol.*, **60**, 723–737.
47. Rapaport,L.R. and Mackie,G.A. (1994) Influence of translational efficiency on the stability of the mRNA for ribosomal protein S20 in *Escherichia coli*. *J. Bacteriol.*, **176**, 992–998.
48. Tock,M.R., Walsh,A.P., Carroll,G. and McDowall,K.J. (2000) The CafA protein required for the 5′-maturation of 16 S rRNA is a 5′-end-dependent ribonuclease that has context-dependent broad sequence specificity. *J. Biol. Chem.*, **275**, 8726–8732.
49. Condon,C. (2006) Shutdown decay of mRNA. *Mol. Microbiol.*, **61**, 573–583.
50. Ruckman,J., Ringquist,S., Brody,E. and Gold,L. (1994) The bacteriophage T4 regB ribonuclease. Stimulation of the purified enzyme by ribosomal protein S1. *J. Biol. Chem.*, **269**, 26655–26662.
51. Uzan,M. and Miller,E.S. (2010) Post-transcriptional control by bacteriophage T4: mRNA decay and inhibition of translation initiation. *Virol. J.*, **7**, 360.
52. Odaert,B., Saida,F., Aliprandi,P., Durand,S., Crechet,J.B., Guerois,R., Laalami,S., Uzan,M. and Bontems,F. (2007) Structural and functional studies of RegB, a new member of a family of sequence-specific ribonucleases involved in mRNA inactivation on the ribosome. *J. Biol. Chem.*, **282**, 2019–2028.
53. Mackie,G.A. (1998) Ribonuclease E is a 5′-end-dependent endonuclease. *Nature*, **395**, 720–723.
54. Jiang,X., Diwa,A. and Belasco,J.G. (2000) Regions of RNase E important for 5′-end-dependent RNA cleavage and autoregulated synthesis. *J. Bacteriol.*, **182**, 2468–2475.
55. Hankins,J.S., Zappavigna,C., Prud'homme-Genereux,A. and Mackie,G.A. (2007) Role of RNA structure and susceptibility to RNase E in regulation of a cold shock mRNA, *cspA* mRNA. *J. Bacteriol.*, **189**, 4353–4358.
56. Arraiano,C.M., Andrade,J.M., Domingues,S., Guinote,I.B., Malecki,M., Matos,R.G., Moreira,R.N., Pobre,V., Reis,F.P., Saramago,M. *et al.* (2010) The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.*, **34**, 883–923.
57. Tedin,K., Moll,I., Grill,S., Resch,A., Graschopf,A., Gualerzi,C.O. and Bläsi,U. (1999) Translation initiation factor 3 antagonizes authentic start codon selection on leaderless mRNAs. *Mol. Microbiol.*, **31**, 67–77.
58. Boni,I.V., Artamonova,V.S., Tzareva,N.V. and Dreyfus,M. (2001) Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. *EMBO J.*, **20**, 4222–4232.
59. Skorski,P., Leroy,P., Fayet,O., Dreyfus,M. and Hermann-Le Denmat,S. (2006) The highly efficient translation initiation region from the *Escherichia coli rpsA* gene lacks a Shine–Dalgarno element. *J. Bacteriol.*, **188**, 6277–6285.
60. Aseev,L.V., Levandovskaya,A.A., Tchufistova,L.S., Scaptsova,N.V. and Boni,I.V. (2008) A new regulatory circuit in ribosomal protein operons: S2-mediated control of the *rpsB-tsf* expression *in vivo*. *RNA*, **14**, 1882–1894.
61. Potapov,A.P. and Subramanian,A.R. (1992) Effect of *E. coli* ribosomal protein S1 on the fidelity of the translational elongation step: reading and misreading of poly(U) and poly(dT). *Biochem. Int.*, **27**, 745–753.
62. Singh,D., Chang,S.J., Lin,P.H., Averina,O.V., Kaberdin,V.R. and Lin-Chao,S. (2009) Regulation of ribonuclease E activity by the L4 ribosomal protein of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **106**, 864–869.
63. Ringquist,S., Jones,T., Snyder,E.E., Gibson,T., Boni,I. and Gold,L. (1995) High-affinity RNA ligands to *Escherichia coli* ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. *Biochemistry*, **34**, 3640–3648.
64. Baker,K.E. and Mackie,G.A. (2003) Ectopic RNase E sites promote bypass of 5′-end-dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.*, **47**, 75–88.
65. Dreyfus,M. (2009) Killer and protective ribosomes. *Prog. Mol. Biol. Transl. Sci.*, **85**, 423–466.
66. Moll,I., Resch,A. and Bläsi,U. (1998) Discrimination of 5′-terminal start codons by translation initiation factor 3 is mediated by ribosomal protein S1. *FEBS Lett.*, **436**, 213–217.
67. Grill,S., Moll,I., Hasenohrl,D., Gualerzi,C.O. and Bläsi,U. (2001) Modulation of ribosomal recruitment to 5′-terminal start codons by translation initiation factors IF2 and IF3. *FEBS Lett.*, **495**, 167–171.
68. Krishnan,K.M., Van,E.W. III and Janssen,G.R. (2010) Proximity of the start codon to a leaderless mRNA's 5′ terminus is a strong positive determinant of ribosome binding and expression in *Escherichia coli*. *J. Bacteriol.*, **192**, 6482–6485.
69. Régnier,P. and Portier,C. (1986) Initiation, attenuation and RNase III processing of transcripts from the *Escherichia coli* operon encoding ribosomal protein

S15 and polynucleotide phosphorylase. *J. Mol. Biol.*, **187**, 23–32.
70. Walz,A., Pirrotta,V. and Ineichen,K. (1976) Lambda repressor regulates the switch between P_R and P_{RM} promoters. *Nature*, **262**, 665–669.
71. Régnier,P. and Hajnsdorf,E. (1991) Decay of mRNA encoding ribosomal protein S15 of *Escherichia coli* is initiated by an RNase E-dependent endonucleolytic cleavage that removes the 3′ stabilizing stem and loop structure. *J. Mol. Biol.*, **217**, 283–292.
72. Hajnsdorf,E. and Régnier,P. (1999) *E. coli* RpsO mRNA decay: RNase E processing at the beginning of the coding sequence stimulates poly(A)-dependent degradation of the mRNA. *J. Mol. Biol.*, **286**, 1033–1043.

SUPPLEMENTARY TABLES

Table S1. Bacterial strains, plasmids and oligonucleotides

Strain	**Relevant characters**[a]	**Reference**
C-1a	*E. coli* C, prototrophic	(1)
C-5686	C-1a *rne-131*~Tn*10*	(2)
C-5699	C-1a *araBp-rpsA*	this work
C-5868	C-1a *rne-3071*~Tn*10*	(2)
C-5869	C-1a rne^{+}~Tn*10*	(2)
C-5874	C-1a Δ*cspE*	this work
C-5899	C-5868 Δ*rng::kan*	this work
C-5901	C-5868 Δ*elaC::kan*	this work
Plasmid		
pGM385	pGM743 derivative, harbours bacteriophage P4 t_{imm}	this work
pGM387	pGM385 derivative, harbours 656303-656798 region of *E. coli* chromosome (*cspE* gene)	this work
pGM396	pGM385 derivative, harbours *E. coli* 3309873-3309373 DNA; the 3309645-3309545 region has been deleted and replaced by the HA DNA fragment (*rpsO*::*HA* construct)	this work
pGM397	pGM396 derivative, harbours λ 37980-37752 (NCBI Accession Number J02459.1) + *E. coli* 3309706-3309373 DNA fragments. The 3309645-3309545 region has been deleted and replaced by the HA DNA fragment (*cI-rpsO* construct).	this work
pGM398	pGM385 derivative, harbours *E. coli* 3309873-3309809+3309706-3309373 DNA; the HA DNA fragment replaces the deleted 3309645-3309545 region (ΔL- *rpsO*::*HA* construct)	this work
pGM743	promoterless pGZ119EH derivative	(3)

pGM924	pGM385 derivative, harbours 656303-656472+656515-656798 region of *E. coli* chromosome (ΔL-*cspE* construct)	this work
pGM928	pGM387 derivative, harbours *E. coli* 656303-656798 DNA with the HA DNA inserted in position 656655 (*cspE*::*HA* construct)	this work
pGM929	pGM924 derivative, harbours *E. coli* 656303-656472+656515-656798 regions with the HA DNA inserted in position 656655 (Δ*L-cspE*::*HA* construct)	this work
pGZ119HE	*ori*V_{ColD}; Cam^R	(4)
pQE31-S1	pQ31 derivative, harbours the *rpsA* gene downstream of *pT5-lacO*	(5)
pREP4	pACYC derivative, harbours the *lacI* gene	Qiagen

Oligonucleotides	**Sequence or Co-ordinates**[b]
HA	CAGCGTAGTCTGGGACGTCGTATGGTTAG
2135	656726-656705
2174	656676-656657
2299	656772-656750
2313	3309646-3309665
2399	656472-656453
2433	CTAATACGACTCACTATAGGTGTCTAAGATTAAAGGTAACG
2434	CTAATACGACTCACTATAGGACACAGCATTTGTGTCTATTTT
2469	3309737-3309756
2521	3309622-3309641

[a] Details about strains and plasmids construction are reported in Materials and Methods section.
[b] Co-ordinates are referred to NCBI Accession Number U00096.2, if not otherwise stated.

Table S2. Ribosomal profile in different conditions of S1 expression.

Fractions	Relative peak area[a]		
	S1 basal[b]	S1 over-expression[b]	S1 depletion[c]
Top	0.34±0.02	0.43±0.03	0.32
30S+50S	0.13±0.02	0.09±0.04	0.14
Monosome	0.13±0.02	0.25±0.01	0.35
Polysome	0.39±0.05	0.15±0.01	0.19

[a] Calculated by weighting the area under the peaks cut out from the ribosomal profile charts (Figure 1A, OD_{260} plots) and normalizing for the weight of the whole profile. Strains and growth conditions are as in Figure 1 legend.

[b] The reported results are the average of two independent determinations.

[c] The results of a representative experiment are shown. In other experiments, the overall profile characterized by a sharp increase of monosome peak was consistently found, but quantitative differences in the relative areas of the peaks from the values reported here were observed, especially for polysome peak area. These variations could be due to fluctuations in the extent of S1 depletion reached at the cell lysis time in different experiments.

Table S3. mRNA expression and stability in the RNase E defective strain

mRNA	R.A.[a]	mRNA half-life (min)[b]	
		rne$^+$	*rne-3071*
cspE$^+$ [c]	1.1	3.5	8.1
Δ*L-cspE*[c]	3.5	< 1	3.9
rpsO$^+$ [d]	17.0	2.0	9.9
Δ*L-rpsO*::*HA*[d]	9.7	0.9	4.9

[a] Cultures of C-5869 (*rne*$^+$) and C-5868 (*rne-3071*) were grown and experiment performed as detailed in Figures 4 and S5 legends and in Materials and Methods. R.A. was calculated as the ratio between mRNA amounts in cultures of mutant and control strains grown up to mid-log phase at 32 °C and further incubated 30 min at 44 °C.

[b] Calculated as detailed in Materials and Methods.

[c, d] Cultures carrying either pGM924 (c) or pGM398 (d). Only signals corresponding to the primary transcripts (from *cspEp* to *cspEt*, c, and from *rpsOp* to *rpsOt*, d) expressed both by the chromosomal (*cspE*$^+$ and *rpsO*$^+$) and plasmid (Δ*L-cspE* and Δ*L-rpsO::HA*) alleles were considered for R.A. and half-life determination.

SUPPLEMENTARY FIGURES

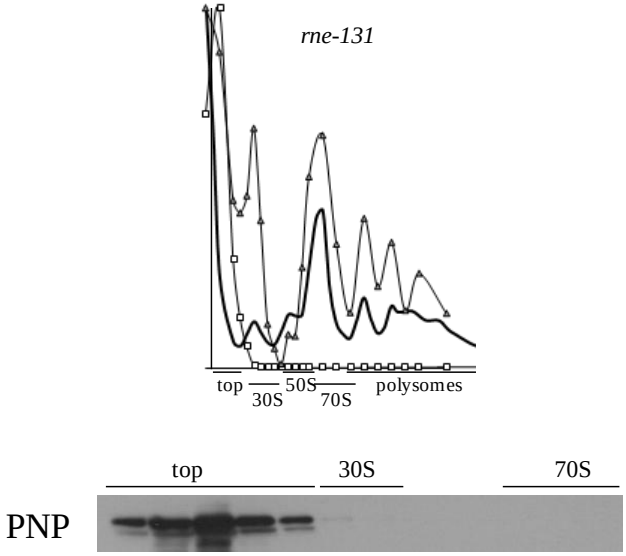


Figure S1. Distribution of PNPase in an RNA degradosome-defective strain. Ribosome profiling and PNPase distribution in an RNA degradosome-defective *rne* mutant. Cultures of C-5686 were grown up to OD600= 0.8 and further incubated 5 min at 37 °C with chloramphenicol (final concentration, 0.1 mg/ml) before the lysis. 14 OD_{260} of crude extracts were fractionated by ultracentrifugation on 10-40% sucrose gradients. Thick continuous line: OD_{260} measured for each gradient fractions; grey triangles: S1 distribution; empty squares, PNPase distribution.

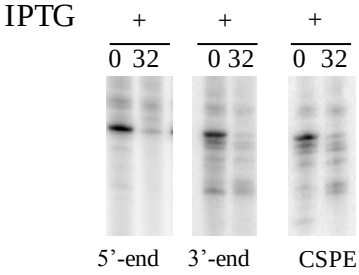


Figure S2. C**haracterization of leaderless *cspE* mRNA ends upon S1 overexpression**. RNA samples from S1-induced C-5874/pREP4/pQE31-S1/pGM924 cultures grown as described in Figure 3 legend, were northern blotted and hybridized with the CSPE riboprobe or with radiolabelled oligonucleotides 2399 and 2135 (5'-end and 3'-end, respectively). RNA was extracted 60 min after S1 induction, immediately before (0) or 32 min (32) following rifampicin-nalidixic acid addition.

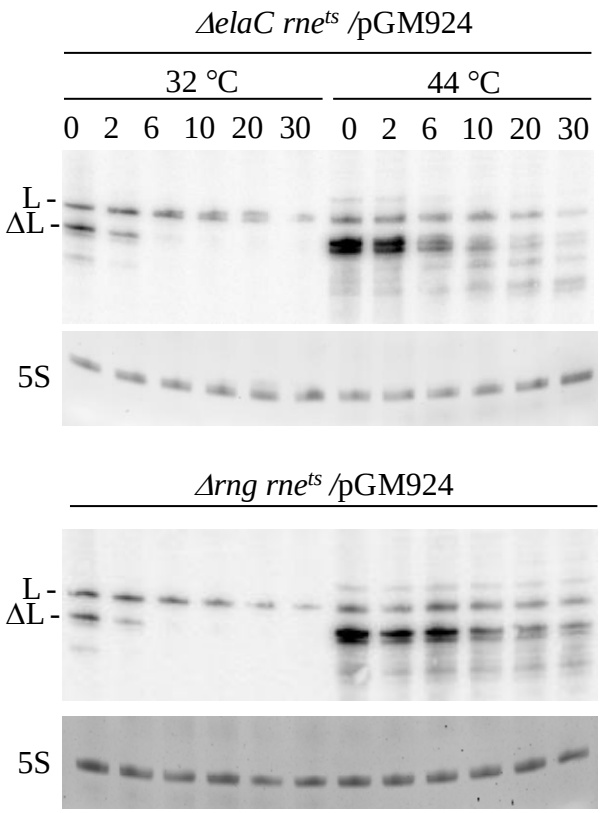


Figure S3. Analysis of RNase Z and RNaseG role in *cspE*$^+$ and Δ*L-cspE* degradation. Cultures of Δ*elaC rne*^{ts} (C-5901) and Δ*rng rne*^{ts} (C-5899) strains carrying pGM924 were grown to mid-log phase at permissive temperature (32 °C) and shifted at non-permissive temperature (44 °C). Rifampicin was added immediately before (32 °C samples) and 30 min after temperature shift (44 °C) and RNA was extracted at the time points indicated on top of the lanes. 5 μg of RNA samples were analyzed by northern blotting with radiolabelled 2135 oligonucleotide. L, leadered *cspE* chromosomal transcript; ΔL, leaderless *cspE* plasmid transcript. Bottom panels: ethidium bromide stained 5S rRNA.

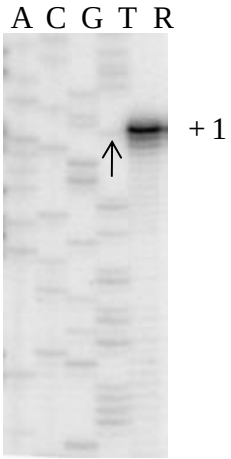


Figure S4. 5'-end mapping of leaderless *rpsO* RNA. RNA was extracted from an exponential culture of C-1a/pREP4/pQE31-S1/pGM398 and analyzed by primer extension with the HA oligonucleotide (R). The position of the 5'-end relative to the first A of the primary transcript (+1) was defined by running the sample along with DNA sequencing reactions obtained with the same oligonucleotide and plasmid pGM398. The arrow on the sequence indicates the signal corresponding to the A of *rpsO* AUG initiation codon. No other signals were detected in the primer extension lane.

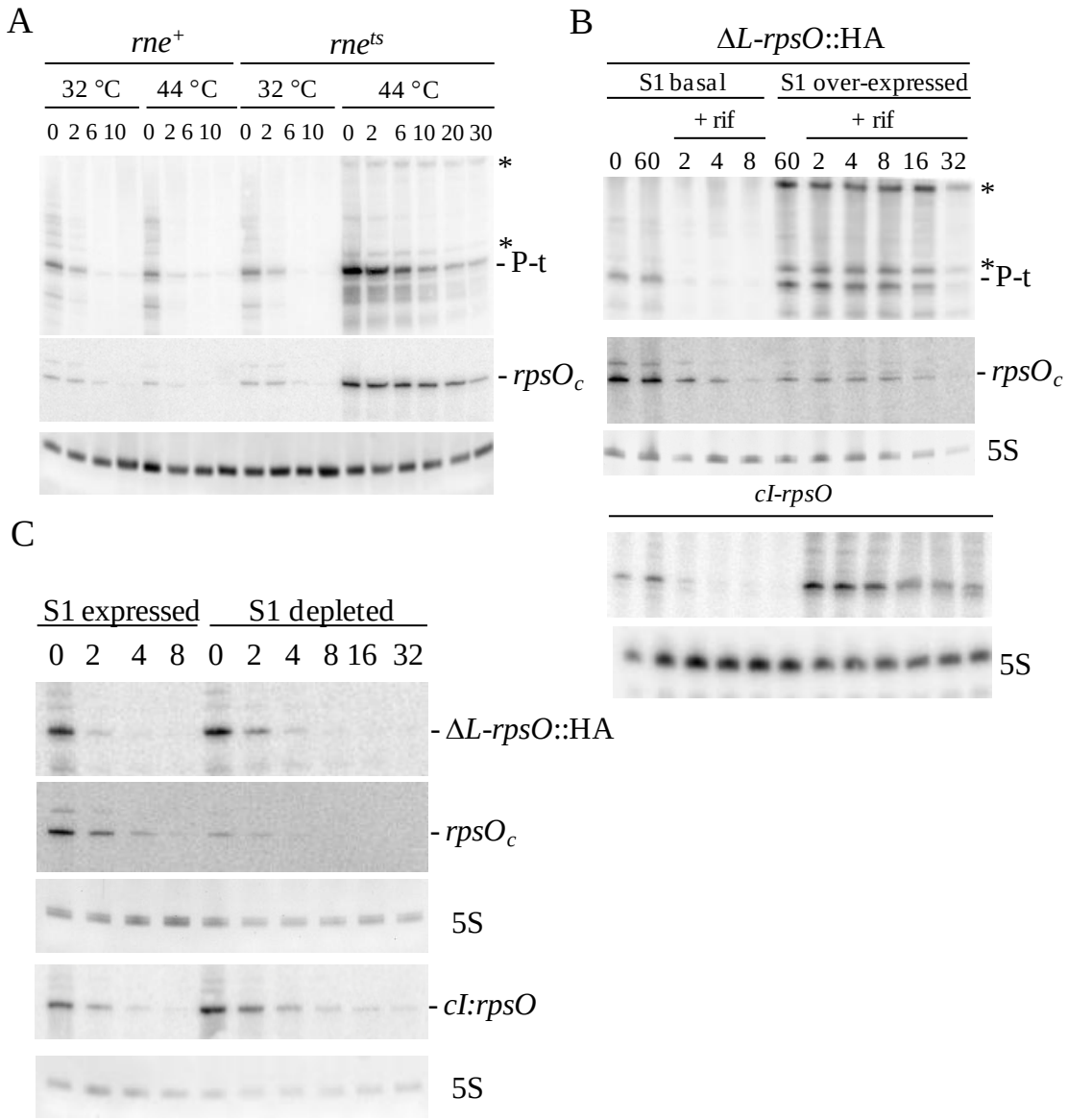


Figure S5. Transcription profile of *rpsO* hybrid alleles. 5 µg of RNA extracted from different strains and conditions were analysed by northern blotting with radio-labelled oligonucleotide probes as detailed below. In all panels 5S rRNA stained with ethidium bromide as a loading control is shown. **A.** Transcription profile of leadered and leaderless *rpsO* mRNA in the RNase E defective strain. Northern blotting was performed with RNA samples extracted from cultures of *rne*$^{+}$ (C-5869) and *rne*^{ts} (C-5868) strains carrying pGM398 (Δ*L-rpsO*::*HA*) grown and processed as described in Figure 4A and hybridized with HA (upper panel) or 2521 (specific for the chromosomal *rpsO* allele; *rpsOc* panel) oligonucleotides. **B**. Expression and stability of leadered and leaderless *rpsO* mRNA upon S1 over-expression. Northern blotting was performed with RNA samples extracted from cultures of C-1a/pREP4/pQE31-S1 carrying pGM398 (Δ*L-rpsO*::*HA*) or pGM397 (*cI-rpsO*) grown and treated as described in Figure 3A and hybridized with radiolabelled HA oligonucleotide. Transcripts size was assessed by running the samples on the same gel with an RNA ladder (RiboRuler™ Low Range RNA Ladder, Fermentas) and by hybridization with proper oligonucleotides (data not shown). P-t, signal corresponding to the *rpsOp-rpsOt* mRNA. **C**. Bacterial cultures of C-5699 carrying pGM398 (Δ*L-rpsO*::*HA* and *rpsOc* panels) or pGM397 (*cI-rpsO* panel) were grown up to OD_{600} = 0.4, diluted 1:4 in permissive (S1 expressed) or non permissive (S1 depleted) conditions and further incubated until OD_{600}= 0.4-0.5 was reached. RNA was extracted from rifampicin- nalidixic acid treated and untreated samples and northern blotted with radiolabelled HA (Δ*L-rpsO*::*HA* and *cI-rpsO* panels) and 2469 (*rpsOc* panel) oligonucleotides.

SUPPLEMENTARY REFERENCES

1. Sasaki,I. and Bertani,G. (1965) Growth abnormalities in Hfr derivatives of *Escherichia coli* strain C. *J.Gen.Microbiol.*, 40, 365-376.
2. Carzaniga,T., Briani,F., Zangrossi,S., Merlino,G., Marchi,P. and Dehò,G. (2009) Autogenous regulation of *Escherichia coli* polynucleotide phosphorylase expression revisited. *J.Bacteriol.*, 191, 1738-1748.
3. Zangrossi,S., Briani,F., Ghisotti,D., Regonesi,M.E., Tortora,P. and Dehò,G. (2000) Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in *Escherichia coli*. *Mol.Microbiol.*, 36, 1470-1480.
4. Lessl,M., Balzer,D., Lurz,R., Waters,V.L., Guiney,D.G. and Lanka,E. (1992) Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region in trans. *J.Bacteriol.*, 174, 2493-2500
5. Sukhodolets,M.V. and Garges,S. (2003) Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry*, 42, 8022-8034.

OPEN ACCESS Freely available online

PLoS one

Comparative Profiling of *Pseudomonas aeruginosa* Strains Reveals Differential Expression of Novel Unique and Conserved Small RNAs

Silvia Ferrara, Margherita Brugnoli, Angela De Bonis, Francesco Righetti, Francesco Delvillani, Gianni Dehò, David Horner, Federica Briani*, Giovanni Bertoni*

Dipartimento di Scienze Biomolecolari e Biotecnologie, Università degli Studi di Milano, Milan, Italy

Abstract

Pseudomonas aeruginosa is a highly adaptable bacterium that thrives in a broad range of ecological niches and can infect multiple hosts as diverse as plants, nematodes and mammals. In humans, it is an important opportunistic pathogen. This wide adaptability correlates with its broad genetic diversity. In this study, we used a deep-sequencing approach to explore the complement of small RNAs (sRNAs) in *P. aeruginosa* as the number of such regulatory molecules previously identified in this organism is relatively low, considering its genome size, phenotypic diversity and adaptability. We have performed a comparative analysis of PAO1 and PA14 strains which share the same host range but differ in virulence, PA14 being considerably more virulent in several model organisms. Altogether, we have identified more than 150 novel candidate sRNAs and validated a third of them by Northern blotting. Interestingly, a number of these novel sRNAs are strain-specific or showed strain-specific expression, strongly suggesting that they could be involved in determining specific phenotypic traits.

Citation: Ferrara S, Brugnoli M, De Bonis A, Righetti F, Delvillani F, et al. (2012) Comparative Profiling of *Pseudomonas aeruginosa* Strains Reveals Differential Expression of Novel Unique and Conserved Small RNAs. PLoS ONE 7(5): e36553. doi:10.1371/journal.pone.0036553

Editor: Pierre Cornelis, Vrije Universiteit Brussel, Belgium

Received January 20, 2012; **Accepted** April 4, 2012; **Published** May 10, 2012

Funding: This work was funded by European Commission (grant NABATIVI, EU-FP7-HEALTH-2007-B contract number 223670). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: giovanni.bertoni@unimi.it (GB); federica.briani@unimi.it (FB)

Introduction

Small RNAs (sRNAs) are widespread in bacteria and play critical regulatory roles in several cellular processes [1–4]. The prototype of a bacterial sRNA is a non-coding RNA 50-300 nucleotides long that acts by imperfect base pairing with *trans*-encoded RNA target(s). sRNA-target interaction may lead to modulation of mRNA translation and/or stability [2,4]. Variations on this theme are also known. For instance, some sRNAs modulate the activity of target proteins or act as mRNAs coding for short proteins. Moreover, there is growing evidence that many sRNAs are *cis*-encoded and transcribed antisense to their target RNA [5]. The target genes of sRNAs-mediated regulation belong to several different functional groups. The prevalent view is that sRNAs might target almost all bacterial cell processes [6]. In pathogenic microbes, several sRNAs have been shown to be involved in host-microbe interactions and in the adaptation to the host environment [6]. In recent years, genome-scale searches have led to a remarkable increase in the number of identified sRNAs in bacteria [2]. In this context, our knowledge of the sRNA complement of *Pseudomonas aeruginosa* seemed limited.

P. aeruginosa is a highly adaptable bacterium which thrives in a broad range of ecological niches. In addition, it can infect multiple hosts as diverse as plants, nematodes and mammals. In humans, it is an important opportunistic pathogen in compromised individuals, such as patients with cystic fibrosis, severe burns and impaired immunity [7,8]. The broad habitat and host ranges of *P. aeruginosa* reflect the large variety of structural, metabolic and virulence functions found in its pangenome (being 6.2–6.9 Mbp the size range of the sequenced strain genomes) [9–12] composed of a high proportion (approximately 90%) of conserved core genes and a rather small accessory genome, found in some strains but not in others, which includes genetic elements supposed to be acquired by horizontal transfer. Accessory genetic elements can confer specific phenotypes that are advantageous under the selective pressure of given habitat or host conditions [10]. Interestingly, a study on the highly virulent strain PA14 has suggested that pathogenicity requires not only virulence factors encoded in the two pathogenicity islands of the accessory genome, but also several core genes [13]. Thus, there seems to be some combinatorial effects between accessory and core functions. In addition, it seems likely that the coordination of the expression of such a panoply of functions is accomplished by regulatory networks based on a large number of regulators. Strikingly, the genome of the archetypal strain PAO1 was found to contain among the highest proportions (9–10%) of regulatory genes as compared to other sequenced bacterial genomes, there being more than 500 genes predicted to encode either transcriptional regulators or two-component regulatory system proteins [11,12,14]. In contrast, only a small number (about 40) of regulatory sRNAs have been reported in *P. aeruginosa* [15] whereas, for example, more than 100 sRNAs have been described in *Escherichia coli* and *Salmonella* [1,3,16], whose genomes are considerably smaller than *P. aeruginosa*.

The apparent low proportion of sRNAs in *P. aeruginosa* could reflect either a real paucity of regulatory sRNAs or the limited

number of genome-wide searches that have been performed in this species [17–19]. In addition, only few of the sRNAs experimentally validated in *P. aeruginosa* have been functionally characterized to date; they have been implicated in carbon catabolite repression (CrcZ) [20], in virulence genes expression control (RsmY,Z) [21–23], or in other functions that can be important for survival in the infected host, such as iron uptake and storage (PrrF1) [24] and *quorum sensing* (PhrS) [25]. Finally, despite the variable degree of virulence shown by different *P. aeruginosa* isolates [13], experimental sRNAs screening has been performed only on PAO1. The identification of genes differentially expressed in virulent *vs.* attenuated strains, irrespective of whether they belong to core or accessory genome, can be a valuable approach for dissecting pathogenicity in this bacterium. This would be particularly true for genes encoding regulatory factors, such as sRNAs, whose expression level may in turn influence the expression of multiple target genes.

In this work we aimed at the systematic identification of sRNAs of *P. aeruginosa* by means of the recently developed "sRNA-Seq" approach, an unbiased high-throughput method for the screening of the entire sRNA complement of any organism based on "next-generation" sequencing technologies [26]. We applied the sRNA-Seq method both to PAO1 and to the highly virulent strain PA14, which differ for the presence of about 112 strain-specific gene clusters (54 PAO1-specific and 58 PA14-specific, including the two PA14 pathogenicity islands PAPI-1 and PAPI-2) [13].

By using this approach, we have identified more than 150 novel candidate sRNAs in *P. aeruginosa*. Interestingly, a relevant number of sRNA hits were strain-specific or showed strain-specific expression, strongly suggesting that they could be involved in determining strain-characteristic phenotypic traits. We probed by Northern blotting 71 candidates and confirmed the expression of 52 new sRNAs, with a validation rate above 73%. Our results expand the panel of *P. aeruginosa* sRNAs resulting from previous surveys and strongly indicate that the degree of sRNAs utilization as regulators is consistent with other bacterial species.

Methods

RNA Isolation and Generation of sRNA-Seq Amplicon Libraries

Total RNA was prepared from 25 ml samples of early stationary phase (A_{600} of about 2.6) cultures of *P. aeruginosa* strains PAO1 [14] and PA14 [27] grown at 37°C in 100 ml of Brain Heart Infusion (BHI) rich medium in 500-ml flasks vigorously shaken (120 rpm). The cells were recovered by centrifugation, resuspended in RNAprotect Cell Reagent (Qiagen) and incubated for 5 min at room temperature, pelleted by centrifugation and stored at −80°C until use. Cells were resuspended in TE–lysozyme (10 mM Tris HCl, 1 mM EDTA, 1 mg/ml lysozyme, pH 7.5), incubated at room temperature for 5 min, and lysed by QIAzol Lysis Reagent (Qiagen). Total RNA was then extracted by the RNeasy Mini Kits (Qiagen) according to the manufacturer's instructions, including RNase-free DNase I in-column treatment and modifications to enrich for small RNAs (<200 nt). The quality of the RNA was assessed by denaturing (8 M urea) 6% polyacrylamide gel electrophoresis (dPAGE).

Size selection of RNA ranging from 20 to 500 nt was performed by fractionating 160 µg of total RNA on preparative dPAGE and cutting the gel slice containing 20 to 500 nt long transcripts. RNA from 20 to 500 nt was electroeluted from gel slices in a Model 422 Electro-Eluter (Biorad). For the preparation of amplicon libraries, the purified 20–500 nt RNA fraction of each strain was first tagged at the 3′-end with linker L1 (Table S1), a 5′-monophosphate oligonucleotide starting with three ribonucleotides followed by a sequence of 20 deoxyribonucleotides and terminally protected with an inverted dT (IDT, Integrated DNA Technologies). The sequence of this hybrid oligonucloeotide does not match any sequence in the *P. aeruginosa* genome and is predicted not to form complex secondary structures. 60 µg of RNA was ligated with 78 µg of L1 in T4 RNA ligase Buffer, 10% DMSO, at 16°C with 90 U of T4 RNA ligase (New England Biolabs). After 16 hrs, additional 90 U of T4 RNA ligase was added and incubation prolonged for 8 hrs. To check ligation efficiency, 0.5–1 µg of RNA from the ligation reaction was probed by Northern blotting with [^{32}P]-labelled -oligos AL1 and PA5SRNA02 (Table S1), which probe L1 and 5S rRNA, respectively. To remove non-ligated L1, the ligation mixtures were run on preparative dPAGE and RNA ranging from 40 to 520 nt was electroeluted from gel slices as described above.

RNase H depletion of tRNA and 5S rRNA was performed as previously described [26] with some modifications. 30 µg of L1-$sRNA_{20–500}$ was annealed to 9 nmol of Oligo Mix, an equimolar mixture of 47 oligonucleotides (Table S1) complementary to the 3′-ends of *P. aeruginosa* tRNAs and 5S rRNA. The RNA-DNA hybrids were then digested with RNase H so as to remove the 3′-L1 tail from the small stable RNAs. Depletion efficiency was checked by Northern blotting with [^{32}P]-labelled oligos AL1 and PA5SRNA02. RNA-L1 was then separated from Oligo Mix, tRNA and 5S rRNA degradation products by preparative dPAGE and electroelution from gel slices as described above.

$cDNA_{20–500}$ was generated from 1 µg of 5S/tRNA-depleted L1-$sRNA_{20–500}$ using the SMARTer^{TM} PCR cDNA Synthesis Kit (Clontech), which combines RNA reverse transcription with cDNA 3′-end SMART tailing activity, according to manufacturer's instruction, except that AL1 oligo (Table S1) was used for reaction priming. cDNA was checked by Southern blotting with [^{32}P]- labelled oligo SmarterII A (Table S1). RNA template was removed by RNase A digestion. To remove free AL1 oligo, the cDNA preparations were run on preparative dPAGE and AL1-cDNA-SMART ranging from 60 to 540 nt (AL1-$cDNA_{20–500}$-SMART) was electroeluted from gel slices as described above.

A $cDNA_{20–500}$-derived amplicon library for 454 pyrosequencing (Roche) was obtained by PCR amplification of AL1-$cDNA_{20–500}$-SMART using Advantage® 2 PCR polymerase (Clontech) with primers (Table S1) tailored for 454-sequencing with Roche Multiplex Identifiers (MID) for "barcoding". In particular, MID42 (TCGATCACGT) and MID47 (TGTGAGTAGT) were used to tag amplicons from PAO1 and PA14, respectively. To remove free primers, the PCR reactions were run on preparative dPAGE and amplicons ranging from 130 to 610 nt were electroeluted from gel slices as described above. Amplicons quality and length distribution was checked by Southern blot with [^{32}P]-labelled SmarterII A oligo and by capillary electrophoresis in

Table 1. nstSGR distribution in PAO1 and PA14.

Loci	nstSGR group	PAO1	PA14
Unique	A	9	–
	B	–	20
Conserved	C	43	–
	D	76	76
	E	–	72
Total	220	128	168

doi:10.1371/journal.pone.0036553.t001

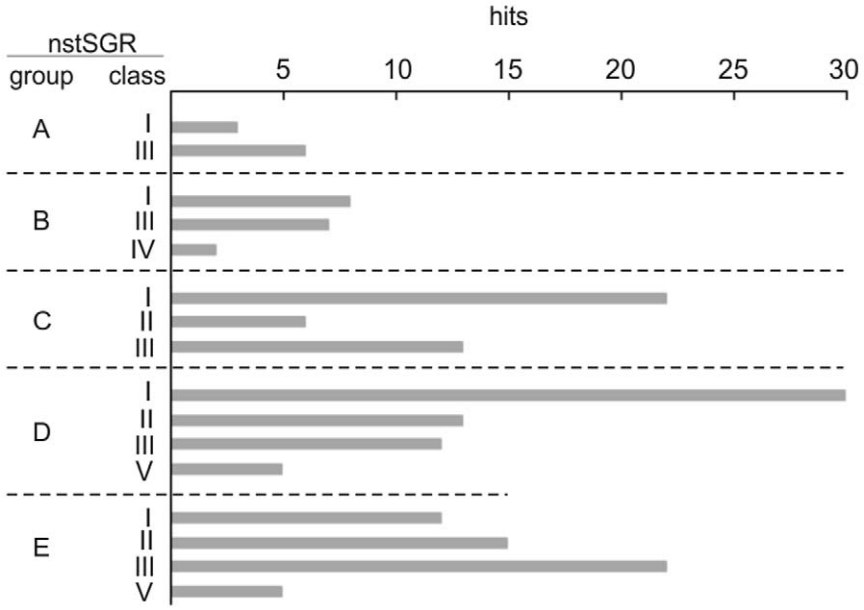


Figure 1. Classes of candidate sRNAs and their distribution within the nstSGR groups resulting from sRNA-Seq. The histogram summarizes the data of Table S2. Candidate sRNAs identified by sRNA-Seq were categorized into five structural/functional classes (I, sRNAs; II, 5′-UTRs; III, asRNAs; IV, CRISPRs; V, sRNAs overlapping annotated ORFs) according to the criteria depicted in Figure S3 and distributed within each nstSGR group (A and B, unique in PAO1 or PA14, respectively; C and E, conserved in both strains but expressed in either PAO1 or PA14, respectively; D, conserved and expressed in both strains).
doi:10.1371/journal.pone.0036553.g001

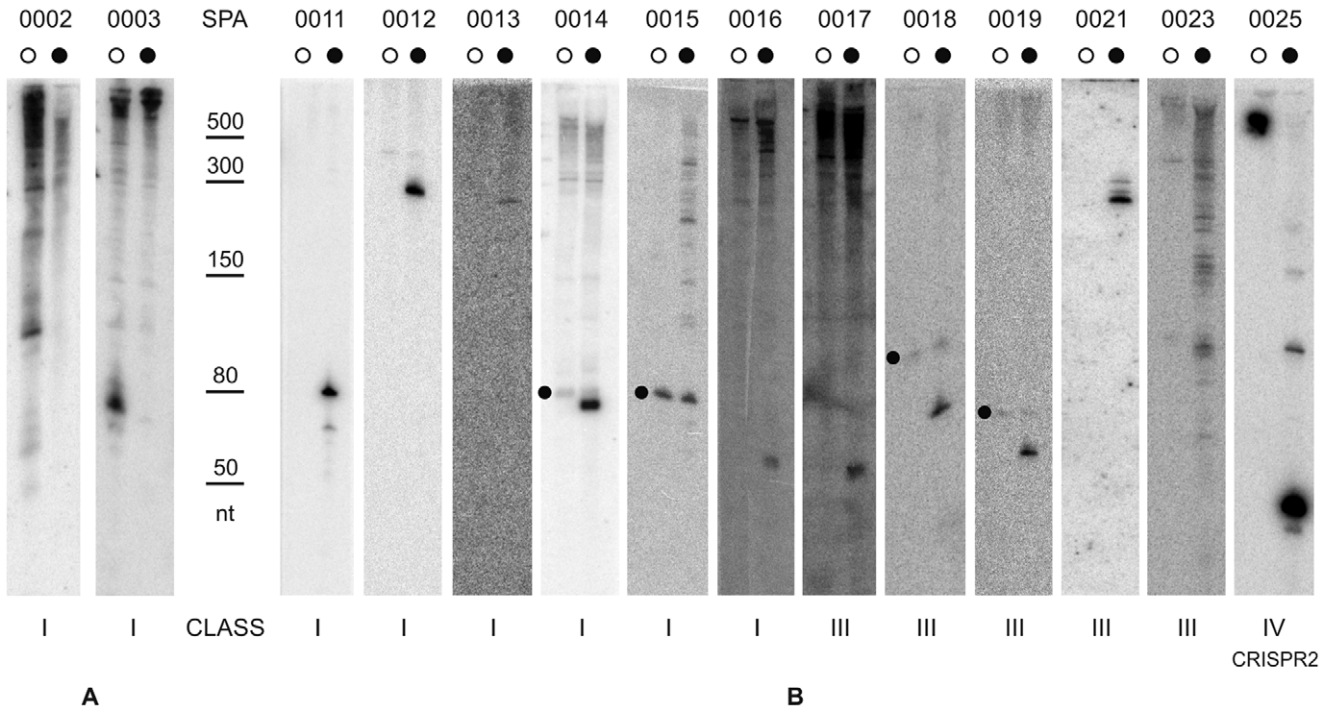


Figure 2. Validation of candidate sRNAs expressed from unique nstSGR in either PAO1 or PA14. A selection of nstSGRs listed in Table S2, unique in either PAO1 or PA14, were inspected by Northern blot for the expression of sRNAs. Total RNA was extracted from both PAO1 (○) and PA14 (●) grown in the same conditions as for sRNA-Seq. Equal amounts of RNA (8 μg) from both strains were blotted and probed with radiolabelled riboprobes (0002 and 0021) or oligos (Table S1) complementary to nstSGR regions with the highest read coverage, as detailed in Materials and Methods. Validated unique sRNAs in PAO1 or PA14 are shown in (A) and (B), respectively. For SPA0014, 0015, 0018, 0019, signals detected in both strains (dots on the left of PAO1 lanes) can be due to aspecific probe hybridization. The ladder of molecular weight markers is shared by (A) and (B). (nt): nucleotides.
doi:10.1371/journal.pone.0036553.g002

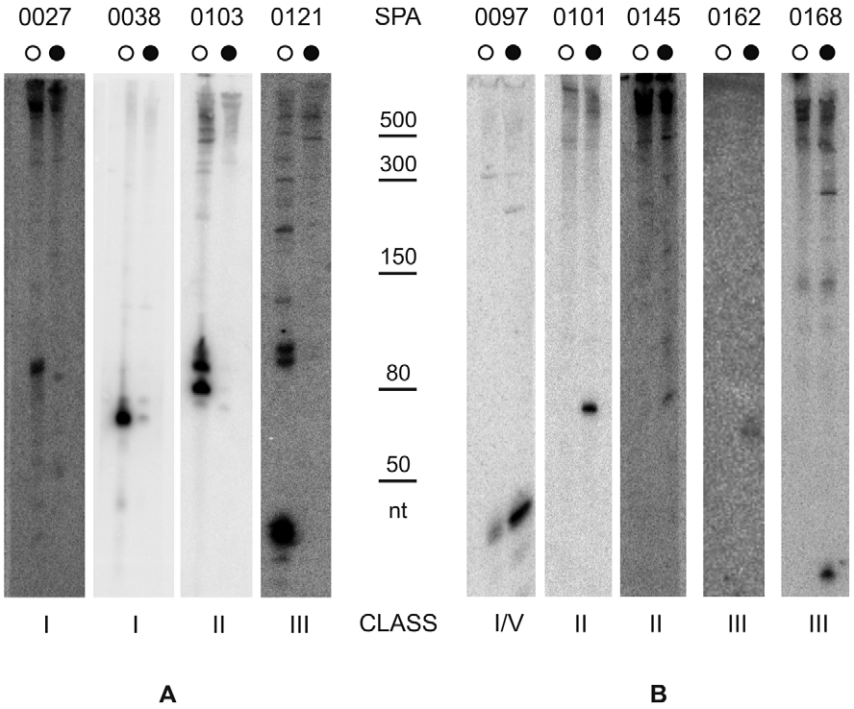


Figure 3. Validation of candidate sRNAs differentially expressed from conserved nstSGR. A selection of conserved nstSGRs listed in Table S2 that were supposed to be differentially expressed between the two strains according to sRNA-Seq data, were inspected by Northern blot. Total RNA was extracted from both PAO1 (○) and PA14 (●) grown in the same conditions as for sRNA-Seq. 8 µg of RNA from both strains were blotted and probed with radiolabelled oligos (Table S1) complementary to nstSGR regions with the highest read coverage. Validated sRNAs which showed higher levels of expression in PAO1 or PA14 are shown in (A) and (B), respectively. The ladder of molecular weight markers is shared by (A) and (B). (nt): nucleotides.
doi:10.1371/journal.pone.0036553.g003

Agilent Bioanalyzer DNA 7500 (Figure S2A). To enrich RNA transcripts ranging from 130 to 500 nt, a second sRNA-derived amplicon library (Figure S2B) was generated from 5S/tRNA-depleted L1-sRNA ranging in size from 150 to 520 nt (selected by preparative dPAGE and electroelution from gel slices) as described above.

Northern and Southern Blot Analyses

The following procedure was used for both Northern and Southern blot analyses. RNA or DNA samples were heated at 95°C for 5 minutes in loading buffer (5 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue dissolved in formamide) and resolved by dPAGE. Nucleic acids were transferred onto Hybond N^+ nylon membranes (GeHealthcare) using a semi-dry electroblotter apparatus (Fastblot B33, Biometra) set at 25 V, 400 mA for 1 hour. The blots were UV-crosslinked and hybridized with [^{32}P]-labelled oligos or riboprobes (Table S1) as described previously [28]. Visualization of radioactive bands was performed by Typhoon^{TM} 8600 variable mode Imager scanner (GE Healthcare BioSciences). All DNA oligonucleotide probes were 5′-end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. Riboprobes were prepared as described previously [28] by T7 RNA polymerase transcription of DNA templates obtained by PCR using oligos listed in Table S1 and *P. aeruginosa* genomic DNA as template. For each validation of candidate sRNA, the probe was complementary to the cognate genomic region with the highest read coverage.

454-pyrosequencing and Data Analysis

Equal amounts of the PAO1 and PA14 amplicon cDNA libraries were combined and submitted to deep-sequencing by a 454 Roche Titanium sequencer using 2/8 of PicoTiterPlate, which should assure at least 140,000 reads. The MID-containing reads were trimmed to eliminate both terminal adaptors, i.e. MIDs, SMART and 454 pyrosequencing primer A-B sequences. Reads were then mapped and clustered throughout the corresponding genome sequence (Genbank accession numbers NC_002516 and NC_008463 for PAO1 and PA14, respectively) as follows. The mapping step was performed using the software SEGEMEHL [29] with default settings but reporting all equal best hits. Mapping positions were considered reliable only if 90% of the read was aligned with ≥90% identity with the genome sequence. Then, the search for read clusters on genome sequences was performed by a sliding window of 200 bp shifted 100 bp at a time along the genome sequence. Significance of clustering of mapped reads was estimated under a null hypothesis of random distribution of reads along the genome using a cumulative Poisson probability. Significant Genomic Regions (SGRs) were defined as consecutive windows where at least one window showed a significant clustering of reads under the Poisson test described above ($P \leq 0.1$). SGRs were divided into "structural" (stSGRs, if the genomic annotation reported the keyword "ribosomal" or "tRNA") and "non-structural" SGRs (nstSGRs, in the other cases). nstSGRs orthology between PAO1 and PA14 strains was determined by reciprocal BLAST. The clustered reads were visualized by GBrowse interface at www.pseudomonas.com database. Sequencing data are accessible at GEO (accession number, GSE36340).

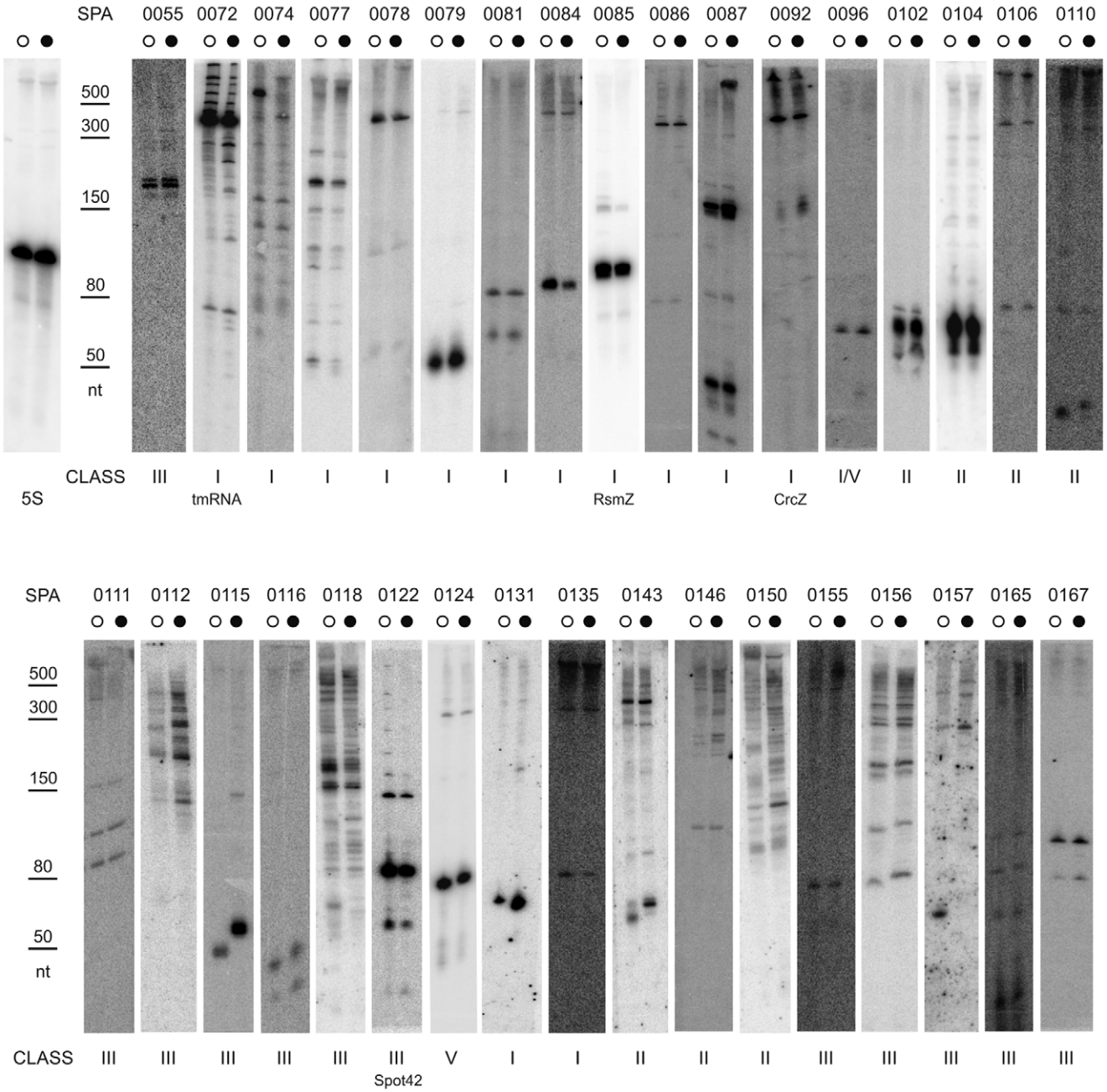


Figure 4. Validation of candidate sRNAs similarly expressed from conserved nstSGR. A selection of conserved nstSGRs listed in Table S2 that were supposed to be similarly expressed between the two strains according to sRNA-Seq data, were inspected by Northern blot. Total RNA was extracted from both PAO1 (○) and PA14 (●) grown in the same conditions as for sRNA-Seq. Equal amounts of RNA (8 µg) from both strains were blotted and probed with radiolabelled oligos or riboprobes (0104, 0112, 0118, 0131, 0143, 0150 and 0157) (Table S1) complementary to nstSGR regions with the highest read coverage. nstSGRs SPA0072, 0085, 0092, and 0122, corresponding to PAO1 *loci* for the known sRNAs tmRNA, RsmZ, CrcZ, and Spot42, respectively, were included in this analysis as positive controls. 5S RNA served as loading control and molecular weight marker. The ladder of molecular weight markers is indicated on the left of each panel. (nt): nucleotides.
doi:10.1371/journal.pone.0036553.g004

Results

Deep-sequencing of the Low Molecular Weight RNA Fraction of *Pseudomonas aeruginosa* PAO1 and PA14 Strains

We aimed at sRNA profiling in the *P. aeruginosa* PAO1 and PA14 strains by sRNA-Seq [26], a massive sequencing approach tailored for unbiased identification of low molecular weight RNA (see Figure S1 for an overview of the procedure). To this end, total RNA was purified from late-exponential cultures of both PAO1 and PA14 strains, respectively, and transcripts ranging from 20 to 500 nt ($sRNA_{20–500}$) were isolated by gel electrophoresis. The 3′ ends of PAO1 and PA14 $sRNA_{20–500}$ were tagged by ligation with linker L1, a mixed ribo-deoxyribo-oligonucleotide with its 3′-end protected by an inverted dT (Table S1), obtaining L1-$sRNA_{20–500}$. sRNA preparations are expected to contain a high proportion of the stable

Table 2. Candidate sRNAs validated by Northern blot.

				PAO1		PA14		
	nstSGR Group[a]	nstSGR name	class	Flanking/Involved *loci*[b]	strand[d]	Flanking/Involved *loci*[b]	strand[d]	Notes[e]
Unique	A	SPA0002	I	2326/2327	−			**MRE**
		SPA0003	I	2729/2730	+			**80**
	B	SPA0011	I			30840/*trbI*	+	20
		SPA0012	I			39480/39500	+	**240**
		SPA0013	I			44640/44650	−	40
		SPA0014	I			49480/49500	+	**80**
		SPA0015	I			60120/60130	−	**MRE**
		SPA0016	I			72510/72520	−	**40**
		SPA0017	III			*trbL*	−	**40**
		SPA0018	III			22270	+	50
		SPA0019	III			35720	+	**50**
		SPA0021	III			59370	+	**240**
		SPA0023	III			59840	−	MRE
		SPA0025	IV			33360	−	MRE;CRISPR-2
Conserved	C	SPA0027	I	*toxR*/0708	+	55150/*toxR*		**90**
		SPA0038	I	2754/*eco*	+	*eco*/28486		**70**
		SPA0055	III	0667	+	08540/*tyrZ*		**SR**
	D	SPA0072[c]	I	*ssrA*	−	53560/53570	+	**MRE**; tmRNA
		SPA0074	I	1429/*lasR*	+	*lasR*/45970	−	**MRE**
		SPA0077	I	*bkdR*/*bkdA1*	+	*bkdA1*/*bkdR*	−	50–70
		SPA0078	I	2421/2422	−	33370/33380	+	20–40
		SPA0079	I	2763/2764	−	28350/28360	+	**50**–70
		SPA0081	I	3069/3070	−	*moxR*/24440	+	**90**
		SPA0084	I	3535/3536	+	18620/18630	−	70–50
		SPA0085[c]	I	*rsmZ*	−	*rpoS*/*fdxA*	+	**120**; RsmZ
		SPA0086	I	3919/1920	−	13170/13190	+	70–**430**
		SPA0087	I	4033/*aqpZ*	+	*aqpZ*/11670	−	**170**
		SPA0092[c]	I	*crcZ*	+	*cbrB*/*pcnB*	+	MRE; CrcZ
		SPA0096	I[f]	2751/2752	−	28520	+	50
		SPA0097	I[f]	2771/2770	+	28250	−	**20**
		SPA0101	II	1244	−	48150	+	**80**
		SPA0102	II	*rpsA*	−	*rpsA*	+	MRE
		SPA0103	II	3229	+	22420	−	**100**
		SPA0104	II	*rhlI*	−	*rhlI*	+	**70**
		SPA0106	II	4133	+	*ccoN*	−	**70**
		SPA0110	II	5473	+	72230	+	**30**
		SPA0111	III	*triC*	−	01970	−	70–140
		SPA0112	III	0367	−	04820	−	**MRE**
		SPA0115	III	2759	+	28410	−	70
		SPA0116	III	2769	+	28290	−	**50**
		SPA0118	III	3350	−	*flgA*	+	**MRE**
		SPA0121	III	5480	−	72350	−	**MRE**
		SPA0122	III	5492	−	*engB*	−	30–**90**; Spot42
		SPA0124	V	1414	+	46160	−	**80**
	E	SPA0131	I	*hasAp*/*hasD*		*hasAP*/*hasD*	−	**60**
		SPA0135	I	2559/2560		31430/31440	+	**80**
		SPA0143	II	*pilU*		*pilU*	+	80

Table 2. Cont.

			PAO1		PA14		
nstSGR Group[a]	nstSGR name	class	Flanking/Involved *loci*[b]	strand[d]	Flanking/Involved *loci*[b]	strand[d]	Notes[e]
	SPA0145	II	*lecB*		*lecB*	−	**90**
	SPA0146	II	3261/3262		21830	+	**120**
	SPA0150	II	*acnA*		*acnA*	+	**140**
	SPA0155	III	*colII*		*colII*	−	MRE
	SPA0156	III	*spuA*		*spuA*	−	**MRE**
	SPA0157	III	*ptsP*		*ptsP*	−	120
	SPA0162	III	*alkB1*		*alkB1*	−	**60**
	SPA0165	III	1735		42100	+	**90**
	SPA0167	III	1166/*pcpS*		49330/*pcpS*	−	80
	SPA0168	III	*purC*		*purC*	+	30

[a]nstSGR group as defined in Table 1: A and B, unique in PAO1 or PA14, respectively; C and E, conserved in both strains but expressed in either PAO1 or PA14, respectively; D, conserved and expressed in both strains.
[b]Name or number (e.g. 2326 is PA2326, 30840 is PA14_30840) of *loci* in the PAO1 and PA14 genomes either overlapping (class II, 5′-UTR; III, asRNA; and V, sense sRNAs overlapping annotated ORFs) or flanking (class I, sRNA) the nstSGRs.
[c]Annotated sRNAs, found by sRNA-Seq, used as a positive control in Northern blot validation experiments.
[d]Upper (+) or lower (−) genomic DNA strand coincident with cDNA reads.
[e]sRNA size predicted by sRNA-Seq. Single values indicates coherent results between PAO1 and PA14. Otherwise, two values (PAO1-PA14) are reported. Values are indicated in bold whenever confirmed by Northern blotting. MRE: Multiple Read Ends defined by non-overlapping reads scattered within the nstSGR. The name of sRNAs used as validation controls is also indicated.
[f]Class assignment in PAO1. The corresponding nstSGR in PA14 was assigned to class V.
doi:10.1371/journal.pone.0036553.t002

and very abundant tRNAs and 5S rRNA that may interfere with the efficiency of sRNA profiling. We thus selectively degraded the stable RNA component as described previously [26]. Briefly, L1-sRNA$_{20–500}$ was mixed with a pool of DNA oligos (Table S1) complementary to *P. aeruginosa* 5S rRNA and tRNAs, and digested with RNase H. Using the SMARTer^{TM} PCR cDNA Synthesis Kit (Clontech), cDNA$_{20–500}$ was then generated from 5S/tRNA-depleted L1-sRNA$_{20–500}$ by reverse transcription with an oligonucleotide primer complementary to L1 (AL1; Table S1) and the cDNA 3′-end was tailed with a specific sequence (see Materials and Methods for details). An amplicon library for 454 pyrosequencing (Roche) was then generated by PCR amplification of cDNA$_{20–500}$ with modular primers complementary to cDNA ends and carrying sequences tailored for 454 sequencing priming and multiplex identification (MID). The PAO1 and PA14 cDNA$_{20–500}$ amplicons described above were combined in a 1:1 ratio (amplicon library 1), and submitted to pyrosequencing. This resulted in a raw pool of 101,019 reads (Figure S2A) among which, 0.3% did not show any identifiable linker sequence. The 100,680 linkers-containing reads were examined for MID sequences. 32,156 and 41,514 reads included MID42 (PAO1) and MID47 (PA14) identifiers, respectively, and were at least 17 bases long. After trimming both terminal linker sequences, the reads showed an average length of 34 and 31 nt for PAO1 and PA14, respectively.

As shown in Figure S2A, sRNA molecules longer than 130 nt were poorly represented in this amplicon library. To increase the abundance of longer RNA molecules (corresponding to a read length of about 230 nt in Figure S2A), additional sRNA-derived amplicons were generated for each strain from 150 to 520 nt long RNA fractionated by gel electrophoresis and processed as described above. These PAO1 and PA14 L1-sRNA$_{150–500}$ amplicons were then combined in a 1:1 ratio, thus producing a second library (amplicon library 2). The pyrosequencing of the latter resulted in a raw pool of 61,490 reads (Figure S2B), among which 59,132 contained identifiable linker sequences. MID analysis showed that 23,608 and 29,107 reads derived from PAO1 and PA14, respectively. Following terminal trimming, the average read length was about 100 nt for PAO1 and 80 nt for PA14.

Identification of Candidate sRNA *Loci* and Comparative Analysis between PAO1 and PA14 Strains

Under stringent mapping criteria (>90% read coverage aligned at >90% identity to reference genome), 13,438 and 22,691 reads gave at least one satisfactory match with the genome sequences of PAO1 and PA14, respectively (GenBank accession numbers NC_002516 and NC_008463). A non uniform distribution of reads across the genomes was observed. In fact, more than 99% of genomic positions showed zero coverage, while a limited proportion of sites showed high levels of coverage. To map candidate sRNA *loci*, genomic regions showing significant reads clustering, hereafter referred to as significant genomic regions (SGRs), were identified as detailed in Materials and Methods. For each strain, about half of the mapped reads clustered in SGRs overlapping stable RNA genes (i.e. tRNAs, 5S rRNA); these were classified as structural SGRs (stSGRs) and not included in further analysis. Around 90% of the remaining mapped reads fell in other significant clusters (non-structural SGRs, nstSGRs), whereas about 10% were not clustered. As stable RNAs are expected to be much more abundant than other RNAs, the observed 1:1 ratio between the number of reads mapping in stSGRs over nstSGR reads indicates the high efficiency of tRNAs and 5S rRNA depletion achieved in amplicon library preparation.

As a whole, we defined 128 and 168 nstSGRs in PAO1 and PA14 genomes, respectively (Table 1 and Table S2) mapping within different *loci*: i) genes for housekeeping RNAs (tmRNA, 6S, 4.5S and RNase P RNAs); ii) genes for sRNAs previously identified

in PAO1 (14) (Table S3), and iii) both intergenic and intragenic *loci* not previously known to express sRNAs (201).

By reciprocal BLAST, we determined whether the identified *loci* were conserved or not in the two strains. Both unique (in either strain) and conserved *loci* were found. Therefore, the corresponding nstSGRs were classified in 5 groups as shown in Table 1. Group A and B nstSGRs map within *loci* unique to PAO1 and PA14, respectively; groups C, D and E include nstSGRs mapping in conserved *loci*. Group D nstSGRs were found in both strains, whereas group C and E nstSGRs were identified only in PAO1 and PA14, respectively. Thus, the comparative profiling of sRNAs from PAO1 and PA14 suggested the existence of both strain-specific (groups A and B) and conserved candidate sRNA *loci*; the latter, in a number of cases (groups C and E), appeared differentially expressed in the two strains.

Within each group described above, we classified the candidate sRNAs according to functional/structural categories established for regulatory RNAs in bacteria [1] as follows (Figure S3). Class I groups nstSGRs located in intergenic regions (>30 nt from flanking ORFs). *Trans*-encoded sRNAs (sRNA) would belong to this class; class II groups nstSGRs with read clustering spanning 5′-untranslated regions (5′-UTRs) in sense orientation. This class would encompass mRNA riboswitches and sRNAs generated by mRNA transcription attenuation or processing; class III includes nstSGRs with intragenic (<30 nt from flanking ORFs) reads clustering in antisense orientation. *Cis*-encoded antisense sRNAs (asRNAs) would cluster in this class; class IV groups intergenic nstSGRs containing CRISPR-like arrays [30]; finally, nstSGRs with read clustering within ORFs and/or 3′-UTRs in sense orientation belong to class V.

The results of this analysis are summarized in Figure 1 and details of each nstSGRs are listed in Table S2. Since class V nstSGRs may correspond to stable mRNA degradation fragments, whose regulatory role is uncertain, they were excluded from further analysis and not reported in Table S2, with the exception of nstSGRs encompassing small putative ORFs.

Remarkably 19 hits of Table S2 corresponded to members of the panel of about 40 sRNAs previously identified in PAO1 [15] including sRNAs annotated in the *Pseudomonas* genome database$_{v2}$ (www.pseudomonas.com) such as the housekeeping tmRNA, 6S, 4.5S and RNase P RNAs, and sRNAs already characterized such as CrcZ, RsmY, RsmZ, PhrS and AmiL [15] and a putative Spot42 sRNA (SPA0122) which is located in a conserved genomic context in *E. coli*, *Salmonella* and pseudomonads [31] (Table S3). We show here that this panel of known sRNAs previously detected in PAO1 is comparably expressed in PA14. Many of the previously identified sRNAs that escaped our analysis have been reported to be expressed at low level or in response to environmental stimuli (e.g. iron limitation for Prrf1 and 2) [24]. However, in a recent deeper transcriptomic survey of PA14 [32] all known *P. aeruginosa* sRNAs were detected. Therefore, it is possible that in our sRNA-seq approach we missed scarcely expressed sRNAs.

Taken together, the data described above, subtracted from those sRNAs already known in *P. aeruginosa*, represent a panel of 163 novel sRNA candidates.

Validation by Northern-blot Analysis of sRNAs Expression from nstSGRs

We tested by Northern blotting the expression of a sample of 71 novel candidates covering all groups and classes (Table S2). Our sample for validation was not random, as we gave priority to strain-specific candidates for validation, but disregarding those having features typical of antisense sRNA regulating a transposase genes (see below). Moreover, we favored class I and III candidates, i.e. *trans*-encoded sRNAs and asRNAs, respectively. In particular, we analyzed 29 class I, 12 class II, 29 class III and 1 class V nstSGRs throughout the A-E groups. The previously identified sRNAs RsmZ (SPA0085) [22], CrcZ (SPA0092) [20], Spot42 (SPA0122) [31] and tmRNA (SPA0072) [19] were used as positive controls. Moreover, a class IV nstSGR, corresponding to CRISPR-2 [33] was included in this validation panel.

Out of 71 novel candidates tested, 52 showed signals in Northern blot experiments (Figures 2, 3 and 4); for 19 we could detect only very faint signals, barely above the background, or no signal (data not shown). Thus, the validation rate was above 73%. Among the validated sRNAs (Table 2), 22 belonged to class I (sRNA), 19 to class III (asRNA), 10 to class II (5′-UTR), 1 to class V. The expression of CRISPR-2 (class IV) was validated and a major band, corresponding to processed crRNA [34], was observed.

The majority of sRNAs tested, which were expected to be equally expressed in both strains (group D) (Table S2), showed signals whose intensity in the two strains was consistent with sRNA-Seq data (i.e. the read number of the corresponding nstSGRs). One exception was SPA0101 (Figure 3B) which showed a comparable read number in both strains, but gave a sharp signal corresponding to a ~70 nt long transcript only in PA14. On the contrary, many group E sRNAs tested (Table S2), whose corresponding nstSGRs displayed expression in PA14 only in sRNA-Seq, showed comparable expression in the two strains in Northern blot analysis. However, all the corresponding nstSGRs of these sRNAs were identified in PA14 by a read number at best slightly above the significance threshold (from 3 to 6). Thus, stochastic fluctuations in amplicon library preparation may have been sufficient to keep the read number below the threshold in PAO1. On the whole, we validated 13 novel unique sRNAs (Figures 2A and B), 30 conserved sRNA with comparable expression in both strains (Figure 4), and 9 conserved sRNAs showing differential expression (Figure 3A and B).

In most cases, we found that transcript size predicted by sRNASeq (Table 2) corresponded to the strongest Northern blotting signal (Figures 2, 3, and 4). The sRNA-Seq reads were scattered within two class I (SPA0015 and SPA0074) and several class II and class III nstSGRs (SPA0023, 0112, 0118, 0150 and 0156). Accordingly, these nstSGRs showed complex patterns with multiple signals of comparable intensity. Degradation of unstable primary transcripts by cellular nucleases may explain these results. However, the presence of scattered reads within an nstSGR was observed also for other sRNAs such as CRISPR2, tmRNA, CrcZ and SPA102, for which a major signal was clearly visible by Northern hybridization. Moreover, in some cases (i.e. SPA0011, 0013, 0018, 0157, 0167, 0168), the regions covered by the reads were smaller than the observed transcripts. As samples preparation for sRNA-Seq included an RNase H digestion step, unspecific RNA degradation by this enzyme may account for these results.

Discussion

Increasing Complexity of *P. aeruginosa* RNA World by sRNA-Seq

In this work we have performed a parallel sRNAs search in *P. aeruginosa* by sRNA-Seq, a powerful unbiased method that allows the analysis by deep sequencing of the whole small transcriptome (*i.e.* both primary and processed transcripts) [26]. Unlike previous surveys performed in PAO1, our search for sRNA *loci* was not biased by *a priori* assumptions about sRNA-based regulation mechanisms, such as binding by Hfq [35], whose role in sRNA-

mediated regulation system is not clearly established in *P. aeruginosa*, or genetic features putatively associated with sRNA-coding *loci* (*e.g.* mapping within intergenic regions with predicted promoters and terminators), which were employed in previous bioinformatics-based analyses [17–19]. The first goal of our analysis was to expand the *P. aeruginosa* sRNA panel resulting from previous surveys in terms of both amplitude and sRNA typologies (potential antisense RNAs for example were completely disregarded by previous analyses) [15,17–19]. Moreover, we did not overlook putative "bifunctional" sRNAs, such as short transcripts encompassing 5′-UTRs or encoding small peptides [3,36].

Our approach resulted in the definition of 163 *loci* expressing new candidate sRNAs. We found a comparable number of class I (sRNA) and III (asRNA) sRNAs, which altogether accounted for more than 75% of our sRNAs panel (Figure 1). In addition, several (34/181) class II sRNAs (mapping within 5′-UTR) were found. These short transcripts, also identified in previous genome-wide searches for sRNAs [37,38], can be generated by premature transcription termination, or 5′-UTR processing as by-products of post-transcriptional gene regulation or mRNA degradation. However, in some cases they can also act as *trans*-encoded sRNAs. In fact, it has recently been reported that two S-adenosylmethionine riboswitches of *Listeria monocytogenes*, SreA and SreB, can base pair with the mRNA of *prfA*, the master regulator of *Listeria* virulence, and repress its expression [39].

Finally, both the annotated CRISPR of PA14 [33] and ten class V *loci* for potential peptide-coding sRNAs were detected (Figure 1). These latter sRNAs may have the dual *status* of short mRNAs (encoding low molecular weight proteins) and *trans*-acting sRNAs, as it has been established for the *E. coli* SgrS and the *Staphylococcus aureus* RNAIII sRNAs [40–42].

We assayed the expression of a large sample of candidate sRNAs by Northern blotting. Remarkably, the expression of many ntsSGRs defined by a read number only slightly above the significance threshold (e.g. SPA0102, 0110, 0112 and 0156) could be demonstrated by Northern blotting. Furthermore, the majority of validated sRNAs showed expression levels in Northern assays that were consistent with sRNA-Seq analysis. Thus sRNA-Seq not only appears a sensitive approach to sRNA identification but could also represent a reliable method for estimating their expression levels in comparative analyses.

On the whole, we could validate the expression of 52 novel sRNAs, more than doubling the number of *P. aeruginosa* sRNAs annotated so far. Interestingly, several validated 5′-UTR nstSGRs (e.g. SPA0101-0104) showed one or two sharp signals in Northern blotting experiments (Figs. 2, 3, and 4) corresponding to discrete RNA species and may thus be good candidates for *trans*-acting sRNAs, as mentioned above. Overall, our data significantly increase the complexity of sRNA complement in *P. aeruginosa* and suggest that RNA-mediated regulation in this organism may be as common and multifaceted as it is in other bacteria [1,3].

sRNA-mediated Regulation May Contribute to *Pseudomonas* Strain-specific Phenotypic Traits

Another purpose of our work was to get hints on sRNA-mediated regulatory mechanisms possibly involved in strain-specific phenotypic traits such as pathogenicity and virulence. To this aim, we performed a comparative analysis of PAO1 and PA14 strains that, although sharing the same host range, differ in virulence, being PA14 considerably more virulent in several model organisms [43].

26 nstSGRs identified by sRNA-Seq consisted of unique *loci* in either PAO1 or PA14 (groups A and B, respectively; Table S2). In PA14, these *loci* mostly mapped within regions of genome plasticity (RGPs, defined as polymorphic strain-specific segments encompassing at least 4 contiguous ORFs) [9], with SPA0016 representing the only exception (Table S2). As for the 9 nstSGRs unique to PAO1, 2 mapped in RGPs (SPA0001, which corresponded to the already known PhrD sRNA [15,19] and SPA0003. Remarkably, 6 overlapped in antisense orientation the 5′-UTRs of a gene encoding a putative transposase of the IS*116*/IS*110*/IS*902* family (SPA0004-8 and SPA0066). This gene is identically repeated six times in PAO1 genome and sRNA-Seq reads were randomly distributed by the mapping software among the six *loci*. Transposase translation regulation by antisense RNAs has been extensively studied in the IS*10* system, where a short RNA (RNA-OUT), which is transcribed in antisense orientation from the 5′-end of the transposase locus, interacts with transposase mRNA to hinder ribosome binding site [44,45]. We did not check by Northern blotting the expression of this PAO1 putative transposase antisense RNA; however, the high overall sRNA-Seq read number (899; Table S2) suggests that it can be actively transcribed and thus may play a role in transposase regulation. Another sRNA antisense to a transposase gene could be expressed by the SPA0022 *locus*, which maps within the PAPI-1 pathogenicity island unique to PA14 and encodes a polypeptide belonging to IS*66* OrfC family [46]. However, in this case only few reads were detected by sRNA-Seq that mapped within the 3′-end region of the transposase.

We validated the expression of 13 novel strain-specific sRNAs by Northern blotting, 2 unique to PAO1 and 11 to PA14. Interestingly, three PA14 novel sRNAs, SPA0015, SPA0021 and SPA0023, fall within the pathogenicity island PAPI-1. SPA0015 *locus* maps in the intergenic region between genes *RL003* (PA14_60130), encoding a homolog of DNA relaxase [47], and *RL004* (PA14_60120; *dcd2*), encoding a putative deoxycytidine deaminase. *RL003* mutants showed virulence-attenuated phenotype [48], and reduced efficiency in PAPI-1 horizontal transfer [47]. SPA0021 and SPA0023 asRNAs are *cis*-encoded antisense to *RL076* (PA14_59370) and *RL033* (PA14_59840) genes, respectively, both encoding hypothetical proteins. An insertion mutation in *RL076* reduced the efficiency of PAPI-1 horizontal transfer [47], whereas an *RL033* mutant showed attenuated virulence [48]. It will be interesting to assess whether sRNA-mediated regulation at these *loci* may be involved in PA14 virulence.

Most nstSGRs (155/181; Table S2) mapped in conserved *loci* and were identified by sRNA-Seq in both strains (group D) or in either one (groups C and E). At the pangenome level, these conserved *loci* mostly belong to the core genome, which constitutes approximately 90% of the total genome and is highly conserved in all strains analyzed so far [9,10]. Only three conserved *loci*, SPA0097, SPA0169 and SPA182 belong to the accessory genome and mapped within RGPs (Table S2). Out of the 52 novel validated sRNAs *loci*, 30 belong to the core genome and exhibited comparable expression in PAO1 and PA14 (Figure 4). On the contrary, 8 sRNAs belonging to the core genome and 1 to RPG43 showed differential expression between the two strains, being more highly expressed either in PAO1 (Figure 3A) or in PA14 (Figure 3B). It has been pointed out that for sRNAs with multiple targets, a hierarchy in target binding due to sRNA-mRNA interaction strength may exist [49]. Thus, it is possible that differential expression of a sRNA may result not only in quantitative differences in the strength of target mRNA(s) regulation, but also ultimately change the number of mRNA species targeted by the same sRNA. It will be interesting to assess whether this is indeed the case for *P. aeruginosa* differentially expressed sRNAs.

Supporting Information

Figure S1 Steps of the comparative analysis of the small transcriptome of PAO1 and PA14 strains. From RNA extraction to sRNAs verification, the sequence of steps followed for the comparative analysis of the small transcriptome of the strains PAO1 and PA14 is depicted as a flow chart. The whole approach was performed in three phases: (A) RNA preparation and 454 pyrosequencing; (B) Deep-sequencing data analysis; (C) Comparative PAO1 *vs* PA14 analysis.
(TIF)

Figure S2 Read length distribution of amplicon libraries. 454 pyrosequencing results in terms of length distribution of untrimmed reads are shown for amplicon library 1 (A) and 2 (B). Terminal adaptors of amplicons for pyrosequencing are altogether 100 nt long. Note the enrichment in (B) of reads with actual length longer than 130 nt, which are scarcely represented in (A).
(TIF)

Figure S3 Criteria for categorization of nstSGRs into classes. nstSGRs are represented by thick black arrows. By way of example, the upper nstSGR bears on top the read cluster by which nstSGR has been defined. Grey tip-ended segments represent annotated ORFs located farther than 30 nt from the nearest end of read cluster. Orange tip-ended segments represent annotated ORFs located within 30 nt from or overlapping with (ol) at least one end of the read cluster. Class I (intergenic), II (5′-UTRs), III (antisense) and V (intragenic) nstSGRs differ for their relative positioning to annotated ORFs. Class IV nstSGRs corresponding to CRISPR-like array are not depicted in this figure.
(TIF)

Table S1 Oligonucleotides.
(PDF)

Table S2 Compilation of nstSGRs identified by parallel sRNA-Seq approach in the *P. aeruginosa* strains PAO1 and PA14.
(PDF)

Table S3 Previously identified *P. aeruginosa* sRNAs found in this work.
(PDF)

Acknowledgments

The authors acknowledge all members of the lab for helpful discussion and technical support, F. Dal Pero and A. Albiero for 454-pyrosequencing and data analysis, respectively, and G. Pavesi and M. Borsani for L1 oligonucleotide design.

Author Contributions

Conceived and designed the experiments: SF FB GB. Performed the experiments: SF MB ADB FR FD. Analyzed the data: SF MB ADB FR FD GD DH FB GB. Wrote the paper: SF GD DH FB GB. Designed the software used in this analysis: DH. Conceived the project: GB.

References

1. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 136: 615–628.
2. Liu JM, Camilli A (2010) A broadening world of bacterial small RNAs. Curr Opin Microbiol 13: 18–23.
3. Gottesman S, Storz G (2010) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol.
4. Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell 43: 880–891.
5. Georg J, Hess WR (2011) *cis*-antisense RNA, another level of gene regulation in bacteria. Microbiol Mol Biol Rev 75: 286–300.
6. Papenfort K, Vogel J (2010) Regulatory RNA in bacterial pathogens. Cell Host Microbe 8: 116–127.
7. Pier GB, Ramphal R (2005) *Pseudomonas aeruginosa*. In: Mandell GL, Bennett JE, Dolin R, eds. Principles and practice of infectious diseases. Philadelphia, PA: Elsevier Churchill Livingstone. pp 2587–2615.
8. Wagner VE, Filiatrault MJ, Picardo KF, Iglewski BH (2008) *Pseudomonas aeruginosa* virulence and pathogenesis issues. In: Cornelis P, ed. *Pseudomonas* genomics and molecular biology. Norfolk: Caister Academic Press. pp 129–158.
9. Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, et al. (2008) Dynamics of *Pseudomonas aeruginosa* genome evolution. Proc Natl Acad Sci U S A 105: 3100–3105.
10. Kung VL, Ozer EA, Hauser AR (2010) The accessory genome of *Pseudomonas aeruginosa*. Microbiol Mol Biol Rev 74: 621–641.
11. Klockgether J, Cramer N, Wiehlmann L, Davenport CF, Tummler B (2011) *Pseudomonas aeruginosa* genomic structure and diversity. Front Microbiol 2: 150.
12. Silby MW, Winstanley C, Godfrey SA, Levy SB, Jackson RW (2011) *Pseudomonas* genomes: diverse and adaptable. FEMS Microbiol Rev 35: 652–680.
13. Lee DG, Urbach JM, Wu G, Liberati NT, Feinbaum RL, et al. (2006) Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. Genome Biol 7: R90.
14. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, et al. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature 406: 959–964.
15. Sonnleitner E, Haas D (2011) Small RNAs as regulators of primary and secondary metabolism in *Pseudomonas* species. Appl Microbiol Biotechnol 91: 63–79.
16. Vogel J (2009) A rough guide to the non-coding RNA world of *Salmonella*. Mol Microbiol 71: 1–11.
17. Livny J, Brencic A, Lory S, Waldor MK (2006) Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. Nucleic Acids Res 34: 3484–3493.
18. Gonzalez N, Heeb S, Valverde C, Kay E, Reimmann C, et al. (2008) Genome-wide search reveals a novel GacA-regulated small RNA in *Pseudomonas species*. BMC Genomics 9: 167.
19. Sonnleitner E, Sorger-Domenigg T, Madej MJ, Findeiss S, Hackermuller J, et al. (2008) Detection of small RNAs in *Pseudomonas aeruginosa* by RNomics and structure-based bioinformatic tools. Microbiology 154: 3175–3187.
20. Sonnleitner E, Abdou L, Haas D (2009) Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 106: 21866–21871.
21. Kay E, Humair B, Denervaud V, Riedel K, Spahr S, et al. (2006) Two GacA-dependent small RNAs modulate the quorum-sensing response in *Pseudomonas aeruginosa*. J Bacteriol 188: 6026–6033.
22. Heurlier K, Williams F, Heeb S, Dormond C, Pessi G, et al. (2004) Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. J Bacteriol 186: 2936–2945.
23. Brencic A, McFarland KA, McManus HR, Castang S, Mogno I, et al. (2009) The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. Mol Microbiol 73: 434–445.
24. Wilderman PJ, Sowa NA, FitzGerald DJ, FitzGerald PC, Gottesman S, et al. (2004) Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. Proc Natl Acad Sci U S A 101: 9792–9797.
25. Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, et al. (2011) The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. Mol Microbiol 80: 868–885.
26. Liu JM, Livny J, Lawrence MS, Kimball MD, Waldor MK, et al. (2009) Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. Nucleic Acids Res 37: e46.
27. Rahme LG, Stevens EJ, Wolfort SF, Shao J, Tompkins RG, et al. (1995) Common virulence factors for bacterial pathogenicity in plants and animals. Science 268: 1899–1902.
28. Briani F, Zangrossi S, Ghisotti D, Deho G (1996) A Rho-dependent transcription termination site regulated by bacteriophage P4 RNA immunity factor. Virology 223: 57–67.
29. Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, et al. (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS Comput Biol 5: e1000502.
30. Sorek R, Kunin V, Hugenholtz P (2008) CRISPR–a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol 6: 181–186.
31. Gottesman S, McCullen CA, Guillier M, Vanderpool CK, Majdalani N, et al. (2006) Small RNA regulators and the bacterial response to stress. Cold Spring Harb Symp Quant Biol 71: 1–11.

32. Dotsch A, Eckweiler D, Schniederjans M, Zimmermann A, Jensen V, et al. (2012) The *Pseudomonas aeruginosa* transcriptome in planktonic cultures and static biofilms using RNA sequencing. PLoS One 7: e31092.
33. Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, et al. (2009) Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. J Bacteriol 191: 210–219.
34. Cady KC, O'Toole GA (2011) Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. J Bacteriol 193: 3433–3445.
35. Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. Nat Rev Microbiol 9: 578–589.
36. Vanderpool CK, Balasubramanian D, Lloyd CR (2011) Dual-function RNA regulators in bacteria. Biochimie 93: 1943–1949.
37. Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jager JG, et al. (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. Nucleic Acids Res 31: 6435–6443.
38. Kawano M, Reynolds AA, Miranda-Rios J, Storz G (2005) Detection of 5′- and 3′-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. Nucleic Acids Res 33: 1040–1050.
39. Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, et al. (2009) A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. Cell 139: 770–779.
40. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J 12: 3967–3975.
41. Wadler CS, Vanderpool CK (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. Proc Natl Acad Sci U S A 104: 20454–20459.
42. Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, et al. (2010) *Staphylococcus aureus* RNAIII binds to two distant regions of coa mRNA to arrest translation and promote mRNA degradation. PLoS Pathog 6: e1000809.
43. Rahme LG, Ausubel FM, Cao H, Drenkard E, Goumnerov BC, et al. (2000) Plants and animals share functionally common bacterial virulence factors. Proc Natl Acad Sci U S A 97: 8815–8821.
44. Simons RW, Kleckner N (1983) Translational control of IS*10* transposition. Cell 34: 683–691.
45. Ma C, Simons RW (1990) The IS*10* antisense RNA blocks ribosome binding at the transposase translation initiation site. EMBO J 9: 1267–1274.
46. Gourbeyre E, Siguier P, Chandler M (2010) Route 66: investigations into the organisation and distribution of the IS*66* family of prokaryotic insertion sequences. Res Microbiol 161: 136–143.
47. Carter MQ, Chen J, Lory S (2010) The *Pseudomonas aeruginosa* pathogenicity island PAPI-1 is transferred via a novel type IV pilus. J Bacteriol 192: 3249–3258.
48. He J, Baldini RL, Deziel E, Saucier M, Zhang Q, et al. (2004) The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. Proc Natl Acad Sci U S A 101: 2530–2535.
49. Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. PLoS Biol 5: e229.

Table S1. Oligonucleotides.

Oligo Name	**Sequence (5' → 3')**
sRNA-Seq amplicon libraries generation[a]	
L1[b]	5'-P-rGrCrUAGTTACTCACACTAGTGTCC/invdT
AL1	GGACACTAGTGTGAGTAACTAGC
SMARTerII A	AAGCAGTGGTATCAACGCAGAGTA
454forAal1MID42	CGTATCGCCTCCCTCGCGCCATCAGTCGATCACGTCTAGTGTGAGTAACTAGC
454revBsmarterIIMID42	CTATGCGCCTTGCCAGCCCGCTCAGTCGATCACGTGTGGTATCAACGCAGAGTA
454forAal1MID47	CGTATCGCCTCCCTCGCGCCATCAGTGTGAGTAGTCTAGTGTGAGTAACTAGC
454revBsmarterIIMID47	CTATGCGCCTTGCCAGCCCGCTCAGTGTGAGTAGTGTGGTATCAACGCAGAGTA
5S rRNA and tRNA depletion	
PA5SRNA01	GAGCTTGACGATGACCTACTCTCACATG
PA5SRNA02	GGAGACCCCACACTACCATCGGCGATG
PAtRNA03	TGGCGGAGAGGGGGGGATTCGAACCCCCGA
PAtRNA05	TGGCGCACTCAGGAGGATTCGAACCTCCGA
PAtRNA06	TGGTGGGTCGTGTAGGATTCGAACCTACGA
PAtRNA07	TGGCGGAGAGATAGGGATTTGAACCCTAGG
PAtRNA08	TGGCGCACCCGGCAGGACTCGAACCTGCGA
PAtRNA09	TGGGGTGGACGATGGGAATCGAACCCACGA
PAtRNA10	TGGTGCGGACGGAGAGACTCGAACTCTCAC
PAtRNA12	TGGCGGAGCGGACGGGACTCGAACCCGCGA
PAtRNA13	TGGCGGTGAGGGAGGGATTCGAACCCTCGA
PAtRNA14	TGGTCGGAGCGACTGGATTCGAACCAGCGA
PAtRNA16	TGGTGCCTCGGGAGAGACTCGAACTCTCAC
PAtRNA17	TGGACGTTCGAGCGGGATTCGAACCCGCGA
PAtRNA18	TGGCAGGGGCGGCTGGATTCGAACCAACGC
PAtRNA19	TGGTGGCTACACCGGGACTTGAACCTGGGA
PAtRNA21	TGGTGCCCAGGAGAAGACTCGAACTTCCAC
PAtRNA22	TGGTGCCCAGGGACGGAATCGAACCGCCGA
PAtRNA23	TGGTGCCGGCACCAGGAGTCGAACCCGGGA
PAtRNA26	GGAAGGCAGTGGGAGTCGAACCCACCC
PAtRNA27	TGGTACCGAGGAGGGGACTCGAACCCCTA
PAtRNA28	TGGTTGCGGGGGCTGGATTTGAACCAACGA
PAtRNA31	TGGCTCCGCGACCTGGACTCGAACCAGGGA
PAtRNA32	TGGTCGGGGTAGAGAGATTCGAACTCCCGA
PAtRNA34	TGGTGCCGGATAGAGGAATCGAACCCCCGA
PAtRNA35	TGGTGGGTCTGGGCAGATTCGAACTGCCGA
PAtRNA36	TGGTGGAGCCAAGGAGGATCGAACTCCTGA
PAtRNA37	TGGTGGAGGGAGAAGGATTCGAACCTTCGA
PAtRNA38	TGGAGCGGGTAGCGGGAATCGAACCCGCA
PAtRNA39	TGGAGCTCATGAGCGGATTTGAACCGCTGA
PAtRNA40	TGGCAGGCCAGGAGGGAATCGAACCCCCAA

Oligo Name	**Sequence (5' → 3')**
PAtRNA41	TGGCGTCCCGGAGAGGGGTCGAACCTCCAA
PAtRNA46	TGGTGGAGCTAGACGGGATCGAACCGTCGA
PAtRNA48	TGGTGGGTGATGACGGGATCGAACCGCCGA
PAtRNA50	TGGCGCAGCGGACGGGACTCGAACCCGCGA
PAtRNA51	TGGCGTCCCCTAGGGGACTCGAACCCCTGT
PAtRNA54	TGGTAGGCACGATTGGATTCGAACCAACGA
PAtRNA55	TGGTCGGGACGGAGTGATTCGAACACTCGA
PAtRNA56	TGGCGGAGGCGGTGAGATTCGAACTCACGG
PAtRNA57	TGGAGCGGGAAACGAGACTCGAACTCGCGA
PAtRNA58	TGGAGGCTGAGGTCGGAATCGAACCGGCGT
PAtRNA59	TGGTGCCCGGAGCCGGGGTCGAACCGGCAC
PAtRNA60	TGGTTGCGGGAGCTGGATTTGAACCAACGA
PAtRNA61	TGGTGGAGCCGGGGGGATTTGAACCCCCGT
PAtRNA62	TGGAGCGGGCGAAGGGAATCGAACCCTCG
PAtRNA63	TGGTGGGCCCACACGGACTCGAACCGTGGA
PAtRNA64	TGGCGCATCCGGCGGGATTCGAACCCACGA
Northern blotting validation probes[c]	
5S	CGCCGATGGTAGTGTGGGGTC
SPA0003	GGCGCTTGAACACCGCTC
SPA0010	ACTGGAACGCCGTCAGGT
SPA0011	TGCCCGAGGCCGGAATCG
SPA0012	CAATTACCGGCGCGGTAGG
SPA0013	GAAGAAGCCCGCAGTAGCG
SPA0014	GGGTGCCGGGAGTTAAGA
SPA0015	TGCGATGCAATTACGCAGTTG
SPA0016	GGTAAGGCCGAGCTGACA
SPA0017	GCGCGTCGGCCGGCGT
SPA0018	AAAATCGGCGAAGCCACTAAAGCACT
SPA0019	CTGGCTAAGGGAGCGGCA
SPA0023	GCCCTCATCAACTCTGCCAAAGAC
SPA0025	TACACGGCAGTGAACACCGCGC
SPA0027	CCGAGCCGATCCCCTACC
SPA0033	AGCCGCAGGTGCAGGAAC
SPA0038	AAAGACCATCTGCGGGGG
SPA0054	CTTGCCGGAGCGGCAAGG
SPA0055	GCAATCCGATCAGAAAGCGC
SPA0056	GGCTGGCGGAGAGCGCTATC
SPA0061	CGGCAACATCGTCGAGAGCGAC
SPA0070	GCCCGGGAGCTCAGCGGA
SPA0071	GTAACCCTGATGGTAGAGCCC
SPA0072	GGTGGAGCCGGGGGGATT
SPA0074	CGCCCGTGCCCGACGAC
SPA0077	CGGTCGTACTGGGTGACG
SPA0078	ACGCGGGCCGCAGGTGGT

Oligo Name	**Sequence (5' → 3')**
SPA0079	ACCCCTCCATGCCCGTCG
SPA0080	GGGAAAGCCCCGGAGGG
SPA0081	AGCAGCCCACCGACCCAG
SPA0084	AGGTGCGGATCTCCGGGG
SPA0085	ATCGTCCTGATGAATCGCCTCCCT
SPA0086	CTGACAGCAGAGGTGAGG
SPA0087	ACATCCCTGTGTCGGAGCA
SPA0088	CCAGCGAGCGCGACATGG
SPA0092	GAAGCTCCCCCCCAAGTAG
SPA0096	GGGACCACAGCGGCAACT
SPA0097	AATCCCGGCCGCGTGGAG
SPA0100	CCAGCCAGGCCGGCGAG
SPA0101	CAGGATCGCGCAGGGTTG
SPA0102	ACGGCGCATTCGTGGACC
SPA0103	GTACCGAGCACTGGCATCC
SPA0106	AGCCGGCGACCGCCGTC
SPA0110	CCCCGCGGGAGTTCGTC
SPA0111	CGCCGCCGTAGCAGAGTC
SPA0113	TGGAACAGGCCGAGCGTGCC
SPA0114	CTGAACTTGCAGCTCATCACTGGG
SPA0115	GAAGAGGGAAGAGCTCCGGCC
SPA0116	AAATCCCGACCGCGGGG
SPA0117	GCGGAGCCCTTGGGCTTG
SPA0119	TCTGCGTAACGTCCAGCTGCAG
SPA0121	ACGCAAGGCCGTCCGGCAC
SPA0122	CCCCGAGCTTCGTATGGG
SPA0124	GGAAAGAGTAGACCGGCGT
SPA0129	CGAGCATGCAACGCGGGA
SPA0135	GCCAAGGCGCAAGCCTGAA
SPA0145	GCCGTCTGGCGCGCGGTA
SPA0146	GGATAACCTTGTGAACAGCCC
SPA0147	ACACACGACTCTCACGGTCG
SPA0155	GGCGCCCACGTGACCCT
SPA0156	GCTGATCATTCCTGCGCTGGG
SPA0160	CGGATGAGGCTTCCATGCTG
SPA0162	GGCGAGGACGGGCGCTA
SPA0163	TTGCTGGTGACCCTGCGC
SPA0165	TGACCTGGTGGCTGCCTG
SPA0167	AAGGTCGAGACAGGACAGTATC
SPA0168	GTTCGGCCTGTTCCATGGCCAG
SPA0174	AGCGCTGAGGCTTGCGAC
DNA templates amplification for riboprobe synthesis[c,d]	
SPA0002F	CCCTACGGGTTACGAGGAGCTC
SPA0002R	**CTAATACGACTCACTATAGGG**GGATCAACACATTCG

Oligo Name	Sequence (5' → 3')
SPA0021F	CTCAAGAGGTTCGGTGTTTCCTCGATTTAAGGG
SPA0021R	**CTAATACGACTCACTATAGGG**CAGGAGCATCAGGCT
SPA0104F	ATCTGGCAGGTTGCCTGCCGTTCATCCTC
SPA0104R	**CTAATACGACTCACTATAGGG**AGTCCCCGTGTCGTG
SPA0112F	GATCGCGGGTCTTGATGCGTGGTGC
SPA0112R	**CTAATACGACTCACTATAGGG**CCGCTTTGTGTACCT
SPA0118F	GCGCCCGTCGAGTCCGCTATTCTGC
SPA0118R	**CTAATACGACTCACTATAGGG**GGCTCGGCTACCTGT
SPA0131F	GGACATTGCGGCAGTCTCC
SPA0131R	**CTAATACGACTCACTATAGGG**TGGCGCTCCTCTCCGGC
SPA0143F	TAGGCGCATTCCTACCTATCCTTGC
SPA0143R	**CTAATACGACTCACTATAGGG**CCTTTTCCACCATCA
SPA0150F	GATCGGGCCACCGCGCATTACC
SPA0150R	**CTAATACGACTCACTATAGGG**AAAGCAGGCATTTCT
SPA0157F	CATGGCGCGGGTCTGGACCAGGAAG
SPA0157R	**CTAATACGACTCACTATAGGG**CGGCGACAAGGCGCT

[a] The MID identifier sequence is underlined.
[b] Ribo-deoxyribo-oligonucleotide with an inverted dT at the 3'-end.
[c] The oligo name corresponds to the cognate nstSGR.
[d] The T7 RNA polymerase promoter sequence is reported in bold.

Table S2. Compilation of nstSGRs identified by parallel sRNA-Seq approach in the *P. aeruginosa* strains PAO1 and PA14.

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
Group A										
SPA0001	I	785400-785600	*phrD*	95	+					PhrD*;* RGP5
SPA0002	I	2568400-2569000	PA2326/PA2327	6	-					
SPA0003	I	3087500-3087800	PA2729/PA2730	36	+					RGP28
SPA0004	III	501100-501300	PA0445	150	+					
SPA0005	III	2556800-2557000	PA2319	149	-					
SPA0006	III	3044700-3044900	PA2690	151	+					
SPA0007	III	3843200-3843400	PA3434	150	+					
SPA0008	III	5383800-5384000	PA4797	150	+					
SPA0066	III	4473500-4473700	PA3993	149	-					
Group B										
SPA0009	I					278200-278400	PA14_03160/PA14_03170	7	+	RGP2
SPA0010	I					1924300-1924500	PA14_22090/PA14_22100	6	-	RGP33
SPA0011	I					2677700-2678200	PA14_30840/*trbI*	11	+	RGP26
SPA0012	I					3515400-3515700	PA14_39480/PA14_39500	152	+	RGP52
SPA0013	I					3974900-3975200	PA14_44640/PA14_44650	5	-	RGP16
SPA0014	I					4401600-4401700	PA14_49480/PA14_49500	8	+	RGP47
SPA0015	I					5355400-5355700	PA14_60120/PA14_60130	10	-	RGP41; PAPI-1
SPA0016	I					6460900-6461400	PA14_72510/PA14_72520	5	-	
SPA0017	III					1315700-1316500	*trbL*	57	-	RGP36
SPA0018	III					1940100-1940300	PA14_22270	7	+	RGP33
SPA0019	III					3176100-3176300	PA14_35720	7	+	RGP23
SPA0020	III					4134300-4134700	PA14_46460	7	+	RGP14
SPA0021	III					5288100-5288500	PA14_59370	9	+	RGP41; PAPI-1
SPA0022	III					5303200-5303500	PA14_59580	5	-	RGP41; PAPI-1
SPA0023	III					5330700-5330900	PA14_59840	5	-	RGP41; PAPI-1
SPA0024	IV					2926400-2927500	PA14_33290/PA14_33300	3	+	RGP24; CRISPR-1
SPA0025	IV					2935700-2937400	PA14_33360	23	-	RGP24; CRISPR-2

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
Group C										
SPA0026	I	580100-580200	*nirS*/*nirQ*	3	-	590077-590177	*nirS*/*nirQ*			
SPA0027	I	781000-781300	*toxR*/PA0708	11	+	4899097-4899397	PA14_55150/*toxR*			
SPA0028	I	912700-912900	PA0836.1	5	-	4738703-4738903	PA53450			P5
SPA0029	I	1028100-1028200	PA0937/PA0938	6	+	4623357-4623457	PA14_52130/*yaiL*			
SPA0030	I	1097000-1097300	PA1013.1/PA1014	5	-	4552631-4552931	PA14_51220/PA14_51230			
SPA0031	I	1135200-1135400	PA1047/PA1048	4	+	4514530-4514730	PA14_50810/PA14_50820			
SPA0032	I	1251200-1251300	*nrdB*/PA1156	5	-	4397712-4397812	*nrdA*/*nrdB*			
SPA0033	I	1474300-1474400	PA1361/PA1362	4	-	4160214-4160311	PA14_46670/*norM*			
SPA0034	I	2031800-2031900	PA1869/PA1870	4	+	3601268-3601369	PA14_40300/PA14_40310			
SPA0035	I	2330400-2330600	*ada*/PA2119	5	-	3312492-3312692	PA14_37170 /*ada*			
SPA0036	I	2558300-2558900	PA2319/*gntR*	2	-	3078836-3079153	*gntR*/PA14_34670			
SPA0037	I	2614800-2614900	PA2364/PA2365	7	+	3029702-3029801	PA14_34070/PA14_34080			
SPA0038	I	3116400-3116600	PA2754/*eco*	11	+	2457373-2457573	*eco*/PA14_28486			
SPA0039	I	3261300-3261400	PA2906/*cobI*	4	-	2307957-2308057	*cobL*/PA14_26485			
SPA0040	I	3489900-3490100	*purF*/PA3109	4	-	2078190-2078390	*cvpA*/*purF*			
SPA0041	I	4094900-4095000	*rpsB*/*map*	3	-	1463019-1463119	*map*/*rpsB*			
SPA0042	I	4780700-4780900	PA4270.1	4	-	745920-746120	*rplL*/*rpoB*			P26
SPA0043	I	5121600-5121800	PA4573/PA4574	4	-	5394065-5394265	PA14_60520/PA14_60530			
SPA0044	I	5325500-5325700	*pnp*/*rps0*	8	-	5597429-5597629	*rps0*/*pnp*			
SPA0045	I	5344900-5345100	PA4758.1	5	-	5616829-5617029	*carA*/*dapB*			P32
SPA0046	I	5439800-5439900	PA4844/*dipZ*	5	+	5710452-5710552	*ctpL*/*dipZ*			
SPA0047	I	6110800-6111000	*aspA*/PA5430	5	-	6383723-6383923	*aspA*/PA14_71670			
SPA0048	II	990900-991100	*rsmA*	4	+	4660491-4660691	*rsmA*			
SPA0049	II	1182800-1183100	*fliC*	5	+	4466831-4467131	*fliC*			
SPA0050	II	3762700-3763000	*flgM*	6	+	1784235-1784535	*flgM*			
SPA0051	II	5337600-5337800	*ftsH*	5	-	5609529-5609729	*ftsH*			
SPA0052	II	5522100-5522300	*azu*	6	-	5792869-5793769	*azu*			
SPA0053	II	5813600-5813800	PA5165	4	+	6085771-6085971	PA14_68230			
SPA0054	III	176100-176200	*pcaG*	15	-	174851-175160	*pcaG*			

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
SPA0055	III	719900-720200	PA0667	8	+	730790-731290	PA14_08540/ *tyrZ*			
SPA0056	III	781400-782200	PA0708	14	+	4898197-4898997	PA14_55150			
SPA0057	III	782900-783600	PA0711	7	+	4896798-4897497	PA14_55117			
SPA0058	III	1027800-1028000	PAO937	4	-	4623557-4623757	*yaiL*			
SPA0059	III	2113600-2114100	PA1933	5	-	3518295-3518795	PA14_39520			
SPA0060	III	3348400-3348700	PA2990	4	-	2220645-2220945	*ugpQ*			
SPA0061	III	3349200-3349400	*sth*	8	+	2219945-2220145	*sth*			
SPA0062	III	3570600-3570800	PA3180	2	-	2003152-2003352	*mutT*			
SPA0063	III	3617500-3617600	PA3230	4	+	1951512-1951612	PA14_22410			
SPA0064	III	3865900-3866200	PA3459	4	-	1674746-1675046	*asnB*			
SPA0065	III	4245900-4246100	PA3788	4	-	1280381-1280581	PA14_15090			
SPA0067	III	5676000-5676200	*pilQ*	4	+	5947414-5947580	*pilQ*			
Group D										
SPA0068	I	332900-333100	PA0295/*spuI*	6	-	347000-347200	PA14_03855/PA14_03860	4	-	
SPA0069	I	586800-587000	*rsmY*	205	+	596800-597000	*rsmY*	99	+	RsmY
SPA0070	I	706700-706900	*vfr*/PA0653	12	-	717700-717800	*vfr*/PA14_08380	4	-	
SPA0071	I	883400-883600	PA0805/PA0806	5	-	4773800-4773900	PA14_53830/PA14_53840	5	+	
SPA0072	I	901500-901900	*ssrA*	1643	-	4749700-4750100	PA14_53560/PA14_53570	1595	+	tmRNA
SPA0073	I	1205000-1205100	PA1112.1	4	-	4444700-4445000	PA14_49990/PA14_50000	7	+	sRNA645
SPA0074	I	1557900-1558200	PA1429/*lasR*	6	+	4086100-4086300	*lasR*/PA14_45970	8	-	
SPA0075	I	1668800-1669200	*ffs*	862		3974500-3974700	PA14_44640/PA14_44650	500	-	4.5S
SPA0076[f]	I					3975600-3975800	PA14_44640/PA14_44650	555	-	4.5S
SPA0077	I	2473000-2473100	*bkdR*/*bkdA1*	8	+	3159700-3159900	*bkdA1*/*bkdR*	8	-	
SPA0078	I	2705600-2705800	PA2421/PA2422	6	-	2937800-2937900	PA14_33370/PA14_33380	7	+	
SPA0079	I	3123200-3123500	PA2763/PA2764	1893	-	2450500-2450700	PA14_28350/PA14_28360	2982	+	
SPA0080	I	3147500-3147600	PA2789/PA2790	7	+	2426600-2426800	PA14_28030/PA14_28040	25	-	
SPA0081	I	3442300-3442500	PA3069/PA3070	5	-	2129000-2129100	*moxR*/PA14_24440	5	+	
SPA0082	I	3705300-3705600	*phrS*	80	-	1841700-1842000	PA14_21260	120	+	PhrS; P20
SPA0083	I	3778000-3778200	*amiL*	28	-	1768500-1768600	PA14_20550/*amiE*	3	+	AmiL
SPA0084	I	3958000-3958200	PA3535/PA3536	14	+	1599900-1600100	PA14_18620/PA14_18630	36	-	

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
SPA0085	I	4057500-4057700	*rsmZ*	3494	-	1500300-1500500	*rpoS*/*fdxA*	8491	+	RsmZ
SPA0086	I	4388800-4389000	PA3919/PA1920	11	-	1135100-1135500	PA14_13170/PA14_13190	7	+	
SPA0087	I	4514400-4514700	PA4033/*aqpZ*	10	+	1011600-1011900	*aqpZ*/PA14_11670	20	-	
SPA0088	I	4519000-4519100	PA4036/PA4037	4	-	1007000-1007300	PA14_11620/PA14_11630	4	+	
SPA0089	I	4536600-4536900	*ribC*/*ribD*	9	-	989600-989800	*ribD*/*ribC*	4	+	sRNA2315
SPA0090	I	4939100-4939300	PA4406.1	10	-	5102700-5102900	*lpxC*/*ftsZ*	3	-	72/101
SPA0091	I	4956300-4956700	*rnpB*	340	-	5119800-5120100	PA14_57460/PA14_57470	630	-	RnpB; sRNA2510
SPA0092	I	5308500-5309000	*crcZ*	100	+	5580500-5581300	*cbrB*/*pcnB*	120	+	CrcZ
SPA0093	I	5372500-5372600	PA4784/PA4785	10	-	5644200-5644400	PA14_63240/*yfcY*	4	-	
SPA0094	I	5884300-5884600	*ssrS*	898	+	6156500-6156800	PA14_69030/PA14_69040	767	+	6S
SPA0095	I(PAO1)/III(PA14)	1997000-1997400	*cysI*/PA1839	8	+	3637200-3637500	PA14_40740	5	-	sRNA1059
SPA0096	I(PAO1)/V(PA14)	3114100-3114200	PA2751/PA2752	11	-	2459700-2460900	PA14_28520	41	+	
SPA0097	I(PAO1)/V(PA14)	3129600-3129700	PA2771/PA2770	15	+	2444300-2444400	PA14_28250	78	-	RGP43
SPA0098	II	410900-411500	PA0367	4	+	425649-426248	PA14_04820	3	+	
SPA0099	II	564700-564900	PA0506	3	+	574700-575000	PA14_06600	6	+	
SPA0100	II	648600-649000	PA0588	9	-	661400-661700	*prkA*	12	-	
SPA0101	II	1348400-1348600	PA1244	8	-	4286000-4286200	PA14_48150	7	+	
SPA0102	II	3549800-3550100	*rpsA*	8	-	2024000-2024200	*rpsA*	5	+	
SPA0103	II	3616900-3617100	PA3229	59	+	1951900-1952200	PA14_22420	3	-	
SPA0104	II	3889700-3899900	*rhlI*	33	-	1651700-1651900	*rhlI*	70	+	
SPA0105	II	3892800-3893300	*rhlA*	6	-	1648400-1649100	*rhlA*	17	+	
SPA0106	II	4622700-4622800	PA4133	11	+	904900-905200	*ccoN*	4	-	
SPA0107	II	5067700-5067900	PA4523	4	-	5231400-5231600	PA14_58690	3	-	
SPA0108	II	5165800-5166000	PA4606	2	-	5438300-5438500	PA14_60960	2	+	
SPA0109	II	5952700-5952900	*glnK*	9	-	6226000- 6226100	*glnK*	13	-	
SPA0110	II	6162000-6162300	PA5473	3	+	6434900-6435100	PA14_72230	6	+	
SPA0111	III	182500-182700	*triC*	12	-	181400-181600	Pa14_01970	7	-	
SPA0112	III	410900-411500	PA0367	2	-	425649-426248	PA14_04820	5	-	
SPA0113	III	449000-449400	*gshB*	9	+	463700-464100	*gshB*	34	+	
SPA0114	III	2982700-2982900	*nuoA*	3	-	2599000-2599200	*nuoA*	11	+	

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
SPA0115	III	3119200-3119700	PA2759	33	+	2454200-2454800	PA14_28410	88	-	
SPA0116	III	3127700-3127900	PA2769	20	+	2446100-2446200	PA14_28290	89	-	
SPA0117	III	3414600-3414800	*pyrD*	5	+	2154400-2154700	*pyrD*	8	-	
SPA0118	III	3761800-3762100	PA3350	14	-	1785200-1785400	*flgA*	5	+	
SPA0119	III	5680700-5681300	*ponA*	27	-	5951200-5952700	*ponA*	11	-	
SPA0120	III	6135900-6136000	*wbpZ*	5	+	6408800-6409000	*wbpZ*	9	+	
SPA0121	III	6171700-6172000	PA5480	14	-	6444600-6445100	PA14_72350	4	-	
SPA0122	III	6183500-6183700	PA5492	36	-	6456400-6456600	*engB*	73	-	
SPA0123	V	784000-784500	PA0713	17	+	4895600-4896200	PA14_55110	12	-	
SPA0124	V	1538600-1539000	PA1414	44	+	4105200-4105600	PA14_46160	197	-	
SPA0125	V	3108800-3109300	PA2747	9	-	2464700-2465200	PA_28600	12	+	
SPA0126	V	3710200-3710700	PA3309	14	+	1836400-1836900	PA14_21220	10	-	
SPA0127	V	5544700-5545000	PA4940	4	-	5816100-5816300	PA14_65260	4	-	
Group E										
SPA0128	I	99747-100047	*fha1*/*tssA1*			97400-97700	PA14_00980/PA14_00990	3	-	
SPA0129	I	517688-518088	PA0458/PA0459			530600-531100	*emrA*/*clpA*	4	+	
SPA0130	I	674110-674210	PA0611/PA0612			686800-686900	*prtR*/PA14_07970	3	+	
SPA0131	I	3813773-3813973	*hasAp*/*hasD*			1725300-1725500	*hasAP*/*hasD*	5	-	
SPA0132	I	3498188-3498388	*fimV*/PA3116			2069900-2070100	PA14_23810/*fimV*	6	+	
SPA0133	I	3221081-3221181	PA2867/PA2868			2348200-2348300	PA14_26990 PA14_27000	3	-	
SPA0134	I	2966810-2967110	*icd*/*idh*			2614800-2615100	*idh*/*icd*	7	-	
SPA0135	I	2893791-2894091	PA2559/PA2560			2733300-2733600	PA14_31430/PA14_31440	3	+	
SPA0136	I	2341596-2341796	PA2127/*cupA1*			3301600-3301800	PA14_37060/PA14_37070	4	+	
SPA0137	I	1720609-1720709	*gltA*/*sdhC*			3922800-3922900	*sdhC*/*gltA*	4	-	
SPA0138	I	1163027-1163327	*braC*			4486600-4486900	*braC*	5	+	
SPA0139	I	5282368-5282768	PA4702/PA4703			5554300-5554700	PA14_62250/PA14_62240	8	+	
SPA0140	II	8500-8700	PA0007			8500-8700	PA14_00080	7	+	
SPA0141	II	40479-40679	PA0039			40500-40700	PA14_00480	4	+	
SPA0142	II	141786-142186	PA0122			139700-140100	PA14_01490	17	+	
SPA0143	II	437564-437864	*pilU*			452300-452600	*pilU*	6	+	

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
SPA0144	II	4507749-4507948	PA4026			1018300-1018500	PA14_11750	3	-	
SPA0145	II	3772836-3773036	*lecB*			1774200-1774400	*lecB*	6	-	
SPA0146	II	3649323-3649523	PA3261/PA3262			1897700-1897000	PA14_21830	4	+	
SPA0147	II	3332328-3332928	*rne*			2236400-2237000	*rne*	10	-	
SPA0148	II	2926236-2926636	*gacA*			2655300-2655700	*gacA*	4	+	
SPA0149	II	2059117-2059417	PA1888			3573000-3573300	PA14_40100	5	+	
SPA0150	II	1703023-1703223	*acnA*			3940300-3940500	*acnA*	5	+	
SPA0151	II	1226131-1226232	PA1134			4423700- 4423800	PA14_49720	4	+	
SPA0152	II	0939003-940203	PA0861			4711400-4711600	PA14_53140	4	-	
SPA0153	II	5755147-5756347	*estA*			6028400-6028600	*estA*	5	-	
SPA0154	II	6106977-6107377	*adhA*			6379900-6380300	*adhA*	8	+	
SPA0155	III	131185-131585	*coIII*			129100-129500	*coIII*	4	-	
SPA0156	III	334675-334975	*spuA*			348800-349100	*spuA*	6	-	
SPA0157	III	380028-380328	*ptsP*			394200-394500	*ptsP*	3	-	
SPA0158	III	4751619-4751918	*bfrA*			774900-775200	*bfrA*	6	-	
SPA0159	III	4005030-4005330	PA3573			1552800-1553100	PA14_18090	11	-	
SPA0160	III	3501188-3500688	*leuB*			2067200-2067500	*leuB*	3	-	
SPA0162	III	2911211-2910811	*alkB1*			2672600-2672800	*alkB1*	3	-	
SPA0163	III	2818447-2818747	PA2500			2808600-2808900	PA14_32330	3	+	
SPA0164	III	1996877-1997277	PA1838/PA1839			3637400 -3637800	PA14_40750	5	+	
SPA0165	III	1878304-1878604	PA1735			3756200-3756500	PA14_42100	4	+	
SPA0166	III	1558629-1558929	*lasR*			4085300 -4085600	*lasR*	10	+	
SPA0167	III	1265312-1264912	PA1166/*pcpS*			4383800-4384000	PA14_49330/*pcpS*	3	-	
SPA0168	III	1096531-1096631	*purC*			4553300-4553400	*purC*	40	+	
SPA0169	III	1068967-1069257	PA0987			4583200-4583500	PA14_51540	5	+	RGP7
SPA0170	III	5001909-5002309	PA4473			5165500-5165900	PA14_58060	4	-	
SPA0171	III	5301669-5301869	PA4722			5573600-5573800	*aspC*	4	+	
SPA0172	III	5868087-5868387	PA5212/*gcvP*			6140300-6140600	PA14_68840/*gcvP1*	7	+	
SPA0173	III	5885587-5885887	PA5229			6157800-6158100	PA14_69050	6	-	
SPA0174	III	5925886-5926066	*argH*			6198100-6198300	*argH*	3	-	

		PAO1				PA14				
nstSGR name[e]	**class**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Genomic coordinates (left-right)**	**Flanking/ Involved *loci*[a]**	**# reads[b]**	**strand[c]**	**Notes[d]**
SPA0175	III	5995982-5996181	PA5325			6268900-6269100	PA14_70300	4	-	
SPA0176	III	6176689-6176989	*ampDH2*			6449600-6449900	PA14_72400	4	+	
SPA0177	III	6218466-6218866	PA5525			6491700-6492100	PA14_72890	11	-	
SPA0178	V	4632556-4632856	PA4141			895000-895300	PA14_10360	15	-	
SPA0179	V	3235659-3236359	PA2883			2333200-2333600	PA14_26780	12	-	
SPA0181	V	3108471-3108771	PA2746/PA2747			2465200-2465500	PA14_28610	5	-	
SPA0182	V	1064311-1064707	PA0981/PA0982			4585500-4585900	PA14_51570	6	-	RGP7
SPA0184	V	5169142-5169442	PA4611			5441600-5441900	PA14_61010	7	-	

[a] *Loci* (locus ID or gene name) either overlapping (class II, III and V) or flanking (class I) the nstSGRs.
[b] Total number of cDNA reads identifying the nstSGR.
[c] Upper (+) or lower (-) genomic DNA strand coincident with cDNA reads.
[d] Names of annotated *P. aeruginosa* sRNAs; Region of Genomic Plasticity (RPG) containing the nstSGR.
[e] Tested sRNAs are in boldface.
[f] In PA14 this nstSGR contains a duplication of *ffs* gene for 4.5S between *loci* PA14_44640 and PA14_44650.

Table S3. Previously identified *P. aeruginosa* sRNAs found in this work.

			PAO1	PA14
nstSGR name	**sRNA name**	**Reference**	**# reads[a]**	**# reads[a]**
Group A				
SPA0001	PhrD	[19]	95	-
Group C				
SPA0028	P5	[17]	5	-
SPA0042	P26	[17]	4	-
SPA0045	P32	[17]	5	-
Group D				
SPA0069	RsmY	[17,21]	205	99
SPA0072	tmRNA	[19]	1643	1595
SPA0073	sRNA645	[18]	4	7
SPA0075	4.5S	[18,19,50]	862	500
SPA0076[b]	4.5S	[18,19,50]	-	555
SPA0082	PhrS; P20	[17,19]	80	120
SPA0083	AmiL	[19]	28	3
SPA0085	RsmZ	[22]	3494	8491
SPA0089	sRNA2315	[18]	9	4
SPA0090	72/101	[19]	10	3
SPA0091	RnpB; sRNA2510	[17-19]	340	630
SPA0092	CrcZ	[20]	100	120
SPA0094	6S	[19,51]	898	767
SPA0095	sRNA1059	[18]	8	5
SPA0122	Spot42	[31]	36	73

[a] Total number of reads identifying nstSGR.
[b] In PA14 this nstSGR contains a duplication of *ffs* gene for 4.5S between *loci* PA14_44640 and PA14_44650.

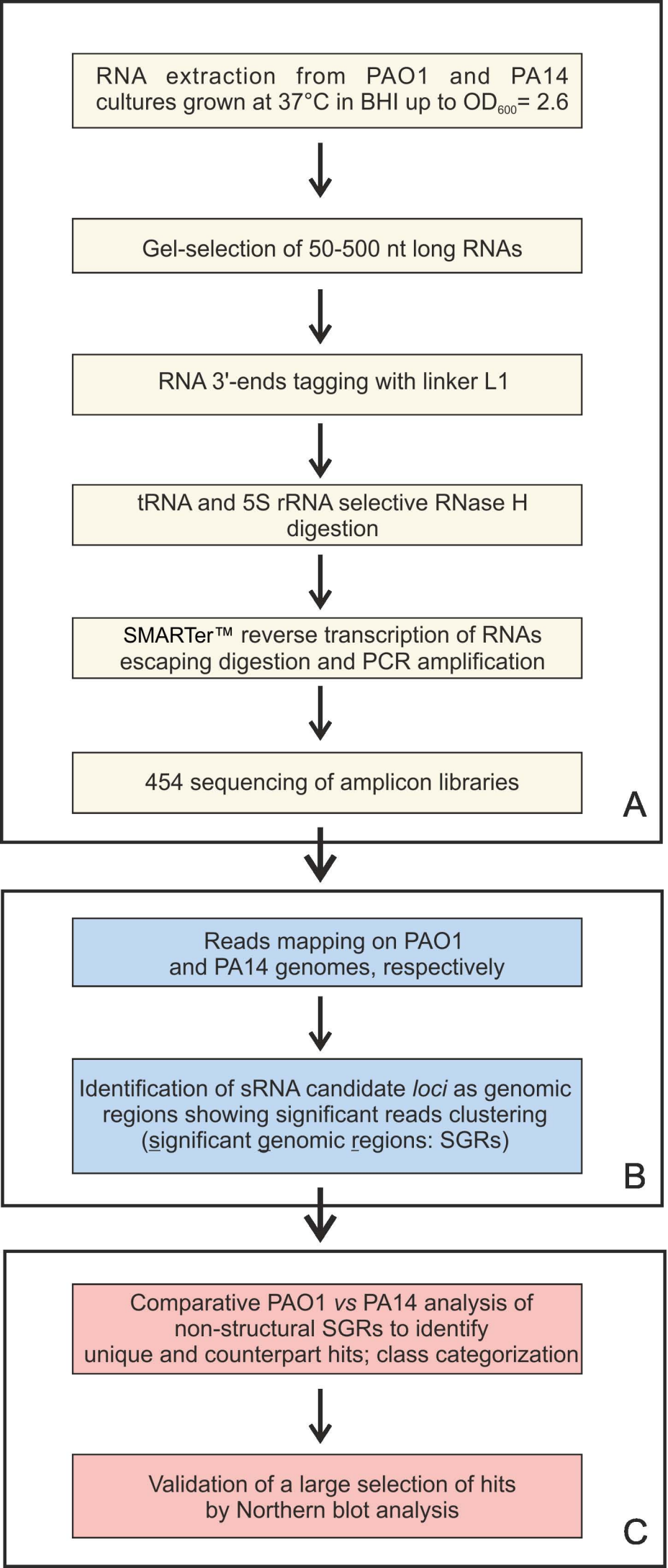
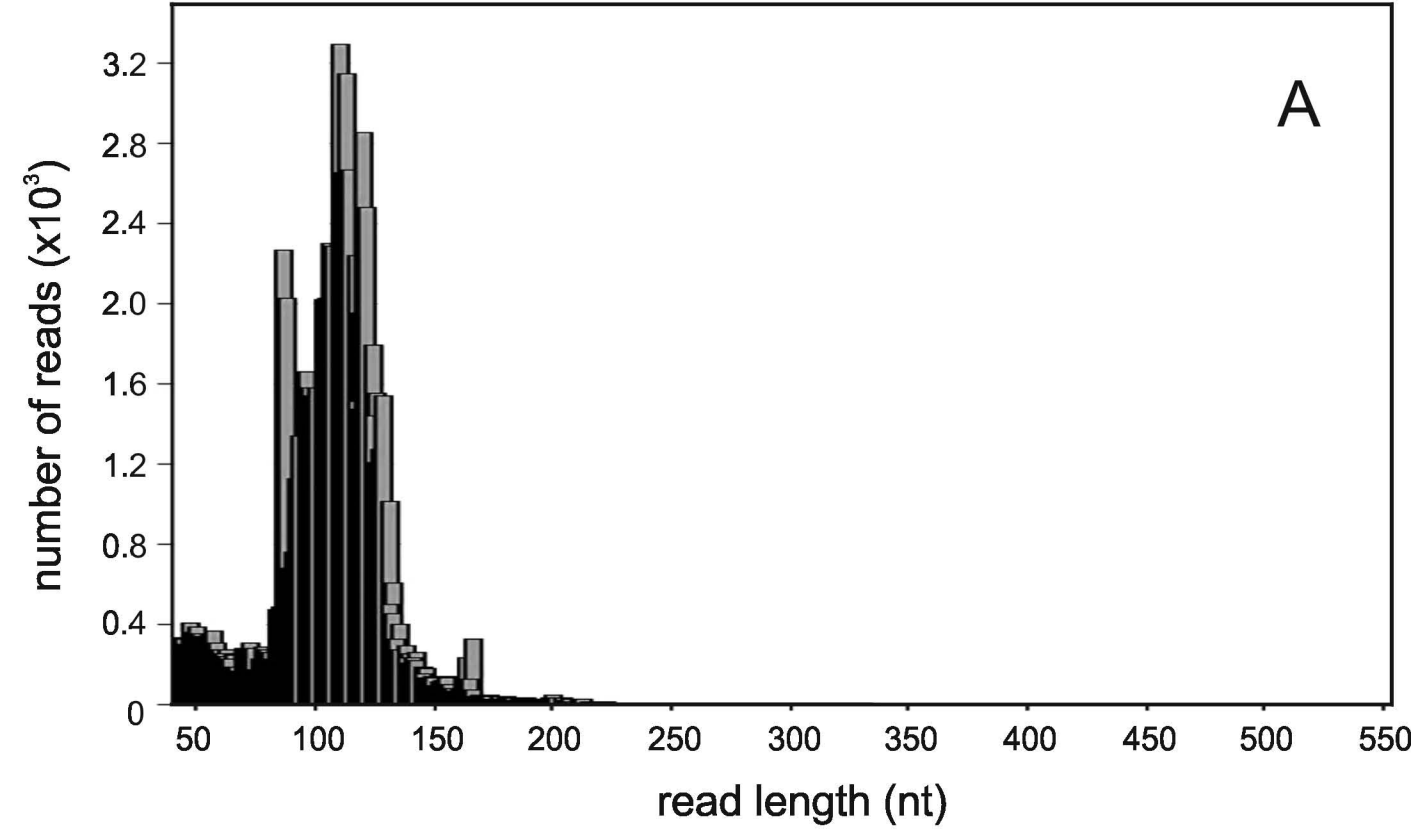

RNA extraction from PAO1 and PA14 cultures grown at 37°C in BHI up to OD_{600}= 2.6
Gel-selection of 50-500 nt long RNAs
RNA 3'-ends tagging with linker L1
tRNA and 5S rRNA selective RNase H digestion
SMARTer™ reverse transcription of RNAs escaping digestion and PCR amplification
454 sequencing of amplicon libraries
A
Reads mapping on PAO1 and PA14 genomes, respectively
Identification of sRNA candidate *loci* as genomic regions showing significant reads clustering (significant genomic regions: SGRs)
B
Comparative PAO1 *vs* PA14 analysis of non-structural SGRs to identify unique and counterpart hits; class categorization
Validation of a large selection of hits by Northern blot analysis
C

Fig. S1


A
number of reads (x10^3)
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
50
100
150
200
250
300
350
400
450
500
550
read length (nt)

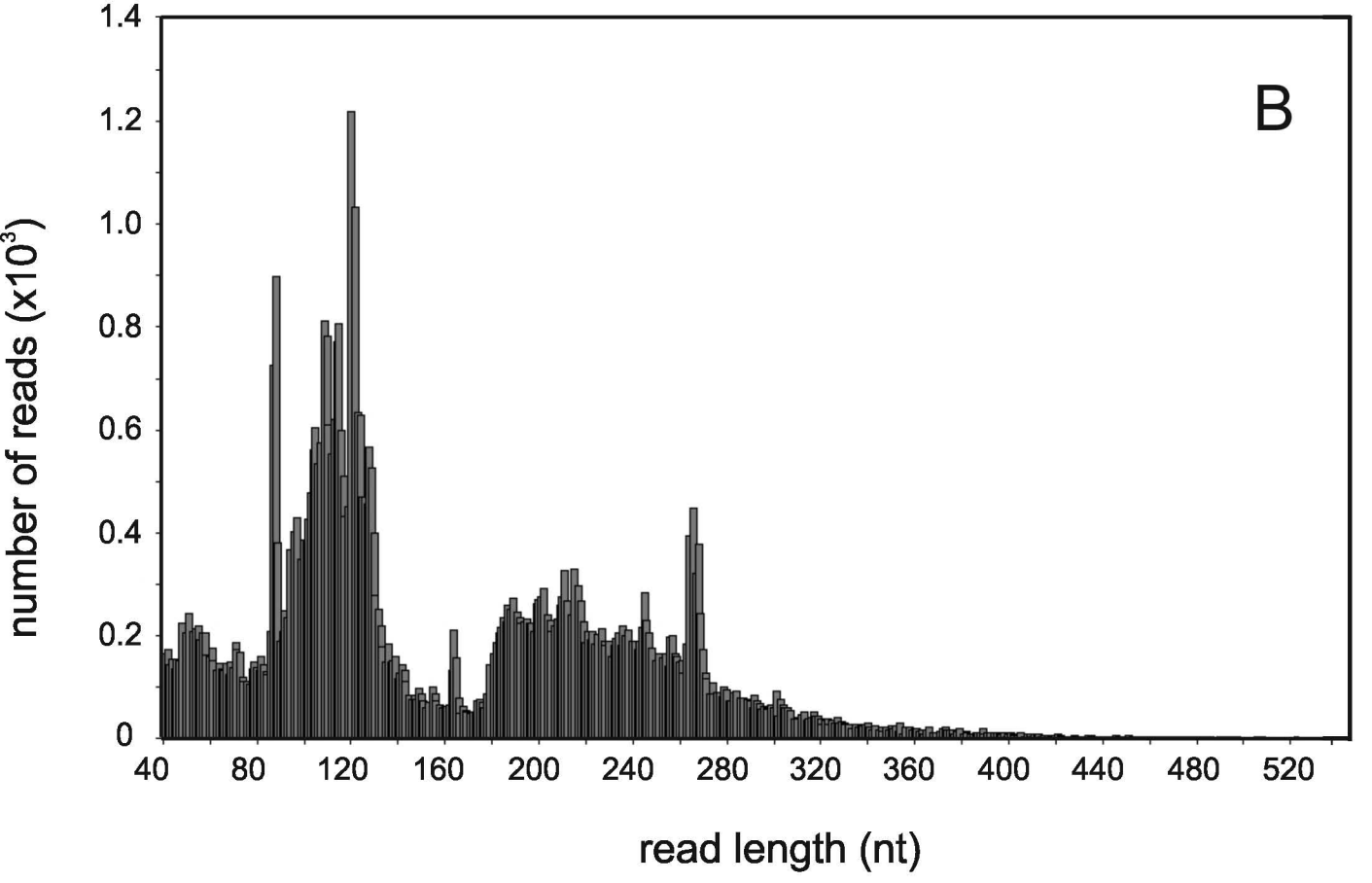

B
number of reads (x10^3)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
40
80
120
160
200
240
280
320
360
400
440
480
520
read length (nt)

Fig. S2

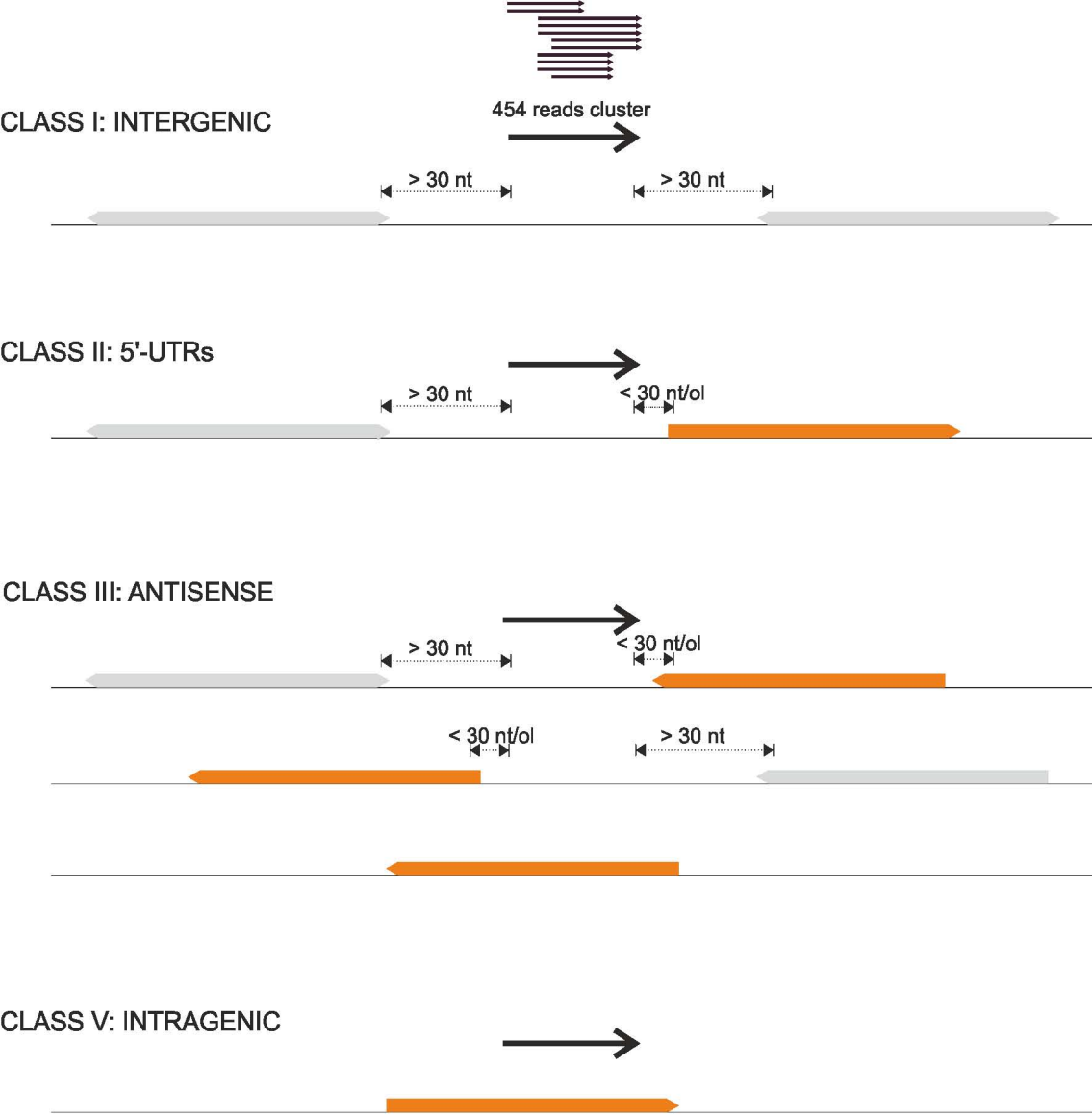

CLASS I: INTERGENIC
454 reads cluster
> 30 nt
> 30 nt
CLASS II: 5'-UTRs
> 30 nt
< 30 nt/ol
CLASS III: ANTISENSE
> 30 nt
< 30 nt/ol
< 30 nt/ol
> 30 nt
CLASS V: INTRAGENIC

FIG. S3

PART III

Content:

Manuscript in preparation: A genetic approach to the identification of *Pseudomonas aeruginosa* RNA thermometers.

A genetic approach to the identification of *Pseudomonas aeruginosa* RNA thermometers

Francesco Delvillani[1] , Clelia Peano[2] , Luca Petiti[2], Christian Berens[3], Christiane Georgi[3], Silvia Ferrara[1], Giovanni Bertoni[1], Gianni Dehò[1], Federica Briani[1*]

[1]Dipartimento di Bioscienze, Università degli Studi di Milano, Italy; [2]Istituto di Tecnologie Biomediche, CNR, Segrate, Italy;[3]Department Biologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany;

*Corresponding author

Dipartimento di Bioscienze

Università degli Studi di Milano, Via Celoria 26, 20133 Milan, Italy

Tel. +39-02-50315033

Fax +39-02-50315044

E-mail: federica.briani@unimi.it

Keywords: RNA regulators in bacteria, post-transcriptional regulation, *dsbA*, *ptxS*

ABSTRACT

Modulation of mRNA translatability either by *trans*-acting factors (proteins or sRNAs) or by in *cis* riboswitches is widespread in Bacteria and controls relevant phenotypic traits. Unfortunately, the identification on a genomic scale of genes post-transcriptionally regulated is not an easy task, as modulation of translation efficiency is not always reflected by changes in the mRNA amount and proteomic analysis has intrinsic technical limitations. We devised a reporter genetic system for the identification of post-transcriptionally regulated genes and we applied this system to search for *Pseudomonas aeruginosa* RNA thermometers, a class of riboswitches that modulate gene translation in response to temperature changes. As *P. aeruginosa* is able to thrive in a broad range of ecological niches, genes differentially expressed at 37 °C *vs.* lower temperatures may be implicated in infection and survival in the human host. We prepared in a plasmid vector a translational fusion library of *P. aeruginosa* DNA fragments (PaDNA) inserted upstream of Tip2, a short peptide able to inactivate the TetR repressor upon expression. The library was assayed in streptomycin resistant merodiploid *rpsL*$^{+}$*/rpsL-31 E. coli* strain in which the dominant *rpsL*$^{+}$ allele (which confers streptomycin sensitivity) was repressed by TetR. PaDNA fragments conferring thermosensitive streptomycin resistance (i.e. putatively expressing PaDNA-Tip2 fusions at 37 °C and not at 28°C) were sequenced. We identified four new putative thermosensors. Two of them were validated with conventional reporter systems in *E. coli*. Interestingly, the two are located upstream of genes previously implicated in *P. aeruginosa* pathogenesis, namely *dsbA* and *ptxS*. *ptxS* RNAT was validated also in *P. aeruginosa* and represents the first RNAT ever described in this bacterium.

INTRODUCTION

Microorganisms can modulate gene expression in response to a variety of chemical and physical signals in order to cope with the challenges posed by a changing environment. Temperature is one of the main physical parameters influencing bacterial growth as it affects both enzymatic reaction rate and macromolecules state. Thus, it is not surprising that bacteria have evolved complex regulatory networks to face sub-lethal temperature changes. In *Escherichia coli*, for instance, both a heat-shock and a cold shock response involving changes in the expression of tens of genes have been described [1, 2]. On the other hand, a modest temperature increase, within the range of permissive growth temperatures for mesophilic bacteria, is used as a signal of the warm-blooded host invasion by some bacterial pathogens to trigger the expression of virulence genes [3, 4].

Either to avoid not amendable cell damages or to establish a successful host infection, the speediness of the response to a sudden temperature variation is critical, thus posing the problem of fast and precise thermosensing. Bacteria exploit different macromolecules as molecular thermosensors [5]; in particular, RNA thermometers (RNATs) have been found to control the expression of a variety of heat shock and virulence genes in Gram negative and Gram positive bacteria (reviewed by [6, 7]). An RNAT can be described as a thermolabile secondary structure that sequesters the translation initiation region (TIR) of the cognate mRNA at low temperature. Local denaturation due to temperature increase allows ribosome binding and mRNA translation. On this very general theme, that apply to most RNATs described so far, variations of the length and localization relative to the AUG of the regions involved in the thermometer structure have been found. For instance, the *E. coli rpoH* thermosensor, for which the definition of RNA thermometer was originally proposed, encompasses RNA regions well within the coding sequence, whereas a widely diffused class of RNATs such as ROSE (Repression Of heat Shock genes Expression) elements form complex secondary structure within the 5'-UTR of the cognate mRNA [8-10]. Moreover, some RNATs, such as those controlling the expression of bacteriophage λ *cIII* or *E. coli cspA*, rely on the temperature-dependent formation of alternative secondary structures, instead of the simple melting of an unstable one [11, 12].

Only two classes of RNATs sharing common structural themes have been defined so far, ROSE and FourU elements; in both cases, sequence conservation within each class is limited to very short stretches of 4-5 bases in the proximity of the Shine-Dalgarno (SD) region [6, 13]. Poor sequence conservation hampers the bioinformatic search of RNATs, which should rely mainly on structural properties; however, the available bioinformatics tools usually overlook non Watson-Crick interactions that seem to play a relevant role in temperature sensing by ROSE elements [14]. Moreover, unique RNATs, completely unrelated to both ROSE and FourU elements, have been

found in Gram positive and Gram negative bacteria making virtually impossible at the moment an exhaustive prediction of RNATs on a genomic scale by the bioinformatics search of conserved structures [4, 8, 11, 15, 16].

We present here a genetic approach, the Tet-Trap, to the identification of RNATs. We have applied the Tet-Trap to *Pseudomonas aeruginosa*, for which RNA thermometers had not reported so far. *P. aeruginosa* is a Gram negative, mesophylic bacterium, endowed with a noteworthy metabolic versatility reflected by a large genome [17]. It can infect hosts as diverse as worms, flies and mammals. In humans it behaves as an opportunistic pathogen and it is responsible of a variety of serious nosocomial infections [18]. As *P. aeruginosa* is a facultative pathogen, it is conceivable that it can exploit the body temperature as a signal for the activation of virulence genes specifically required during the warm-blooded host infection. In facts, in a recent transcriptomic survey by RNA-Seq of *P. aeruginosa* grown at 28 and 37 °C, Wurtzel et al. detected genes preferentially expressed at the body temperature of the mammalian host, among which virulence genes were significantly enriched [19]. By Tet-Trap we could identify four genes post-transcriptionally regulated by a temperature upshift from 28 to 37 °C. Interestingly, two of them, namely *dsbA* and *ptxS*, encode proteins previously involved in *P. aeruginosa* virulence [20, 21].

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains, plasmids and oligonucleotides sequences are reported in Supplementary table S1. *P. aeruginosa* strains used here were PAO1 [17] and PA14 [22, 23] (Genbank accession numbers NC_002516 and NC_008463, respectively). *E. coli* coordinates throughout this work refer to Genbank Accession Number U00096.2.

Construction of E. coli Tet-Trap reporter strains. The SpcR and StrS reporter cassettes were prepared as follows. A cassette constituted by *Tn10* tet regulatory region (*tetRp-tetO-tetAp*) [24] inserted between *kan*^{R} (under *tetAp* promoter) and *lacZ*α (under *tetRp*) genes was integrated into C-1a [25] chromosome between *tonB* and *yciA* loci. The cassette was obtained by three steps-PCR as follows. Three partially overlapping fragments corresponding to *Tn10* tet regulatory region, *lacZ*α and *kan*^{R} genes were synthesized by PCR (oligonucleotide 2602-2603 on C-5868 genomic DNA[26] were used for *Tn10* tet amplification; 2604-2605 on pUC19 [27] for *lacZ*α; 2600-2601 on pQE31S1[28] for *kan*^{R}). Then *Tn10* tet and *lacZ*α fragments were used as templates in a PCR reaction with oligonucleotides 2603-2604, obtaining fragment *lacZ*α-tet. Finally, the full length cassette was obtained by amplification of *lacZ*α-tet and *kan*^{R} fragments with primers 2604-2606 and was cloned in the *Sma*I site of pGM742 [29], giving plasmid pGM932. The insert was excised by digestion with *Not*I and *Xmn*I and integrated in C-1a/pKD46 between nucleotides 1309870 and 1309872 by λ Red-mediated homologous recombination [30], obtaining C-5898. To replace the *kan*^{R} gene with the *aadA::*GFP coding region, conferring spectinomycin resistance and fluorescence, the *aadA::*GFP ORF was amplified from plasmid pZR80-2 [31] with oligonucleotides 2683-2684, harboring 50 nt long tails homologous to the regions flanking *kan*^{R} ORF in C-5898. C-5898/pKD46 was transformed with the above PCR fragment; the recombinants were selected on spectinomycin plates and their fluorescence was evaluated by Versadoc imaging of the plates. The cassette harbored by an highly fluorescent clone was sequenced and revealed a spontaneous A to G transition within *tetA* Shine-Dalgarno region (position 916 of *Tn10*; GenBank accession number AY528506.1). This mutant strain was named C-5899. To replace the *aadA::*GFP gene with *rpsL*, the *rpsL* ORF and the *cat* gene with two flanking FRT sites were amplified by PCR with oligonucleotides 2713-2714 on C-1a genomic DNA and oligonucleotides 2636-2712 on pKD3[30], respectively. The two above partially overlapping amplicons were used as DNA templates for a PCR reaction with oligonucleotides 2638-2602. The resulting DNA fragment was used to transform C-5899/pKD46, obtaining C-5912. The reporter construct was then transduced into C-5708, a C-1a derivative carrying the recessive *rpsL-31* allele (K42T substitution in S12), which confers streptomycin

resistance [32]. The *cat* cassette was excised from the StrS recombinant strain by FLP-mediated recombination, obtaining C-5918. *Pcat-tetR*:*kan*^{R} cassettes with either the *wt* or the mutated -10^{CATTTA} variant of *Pcat* promoter were amplified from WH1001 derivatives [33] with oligonucleotides 2685-2686. Region 806595-808520 of BW25113 genome was replaced with the two cassettes by λ Red-mediated homologous recombination, obtaining strains KG264 and KG265, respectively. We transduced the *Pcat*$^+$*-tetR:kan*^{R} region from KG264 into C-5899 obtaining C-5901; C-5907 is a C-5901 derivative in which the *kan*^{R} cassette was removed by FLP-mediated recombination [30]. C-5920 was similarly obtained by P1 transduction from KG265 into C-5918 and excision of *kan*^{R} gene.

Tet-Trap plasmids. pGM956 and pGM957 carry a chimeric gene composed by an $(SG_4)_5$ linker and *trxA* fused upstream of TIP2. The construct was obtained by three-step PCR. $(SG_4)_5$ was obtained by annealing of the partially overlapping 2689-2690 oligonucleotides and extension with *Pfu* polymerase (Stratagene). *trxA*-TIP2 sequence was amplified by PCR on pWH2354 [33] with oligonucleotides 2691-2692. The final PCR was performed on the two above fragments with oligonucleotides 2692-2693 obtaining the full length construct $(SG_4)_5$-*trxA*-TIP2. This was digested with *Sph*I-*Eco*RI and cloned in pGZ119HE [34] obtaining pGM957. pGM956 is a pGM957 derivative that carries a translation initiation region (TIR) in frame with $(SG_4)_5$-*trxA*-TIP2. The TIR was obtained by annealing oligonucleotides 2617-2618 and cloning the fragment in pGM957 between *Hind*III and *Sph*I sites. Both plasmids were checked by sequencing.

BgaB and EGFP reporter plasmids. To construct the shuttle vector pGM931, the *Hind*III-*Pst*I pGM362 fragment carrying the transcriptional terminator t_Ω [35] was cloned in pBAD24-Δ1 [36], obtaining pGM930. The *Mlu*I-*Hind*III pGM930 fragment carrying araBp-t_Ω was cloned in pHERD20T [37], obtaining pGM931. All reporter constructs were assembled in pGM931 and checked by sequencing. *bgaB* was amplified by PCR from pBAD2_bgaB [38] with primers 2846-2847, digested with *Nco*I and *Pst*I and cloned in pGM931, obtaining pGM978. pGM978 was digested with *Nco*I and *Eco*RI and used as a backbone for translation fusions. The following portions of *ptxS* gene were amplified from *P. aeruginosa* PAO1 genome and cloned: region 2487532-2488013 (primers 2850-2851; plasmid pGM980); 2487779-2488013 region (primers 2851-2852; plasmid pGM981). EGFP ORF was amplified from pZR80-2 *aadA::*GFP with oligonucleotides 2803-2804, digested with *Pst*I-*Kpn*I and cloned in pGM931, obtaining pGM963. Translational fusions were set up by cloning in the *Kpn*I site PAO1 genomic DNA fragments overlapping the *dsbA* locus. The following regions were amplified, digested with *Kpn*I and cloned in pGM963: 6318732-6318445 (oligonucleotides 2806-2808), 6318624-6318445 (2807-2808) and 6318624-6318507(2807-2809) obtaining plasmids pGM964, pGM965 and pGM966, respectively. Control plasmids pGM989 and

pGM991 carry the leader region and first 9 codons of *E. coli recA* in frame with *bgaB* and EGFP genes, respectively. *recA* fragments were amplified from MG1655 genomic DNA with oligonucleotides 2915-2916 (amplicon Rec1) and 2928-2929 (Rec2). Rec1 was digested with *Nco*I and *Eco*RI and cloned in pGM978, obtaining pGM989, whereas PCR2 was digested with *Kpn*I and cloned in pGM963, obtaining pGM991. All pGM931 derivatives were constructed in *E. coli* and transferred in PAO1 by triparental conjugation [39].

Bacterial cultures were grown in LD (10g/l Tryptone, 5 g/l Yeast Extract, 5 g/l NaCl). When needed, media were supplemented as follows: 100 µg/ml ampicillin; 30 µg/ml chloramphenicol; 100 µg/ml spectinomycin; 0.5 mM IPTG; 0.1% arabinose. *P. aeruginosa* cultures carrying pGM931 derivatives were grown in LD supplemented with 300 µg/ml carbenicillin.

Library generation

Genomic DNA was extracted from stationary phase cultures of PAO1 and PA14 with Puregene Kit. 1 µg of PAO1 and PA14 genomic DNA was partially digested with *Alu*I, *Hae*III or the two enzymes for 30 min at 37°. The digestions were loaded on 1% agarose gel and the bands corresponding to DNA ranging from 300 to 800 nt were cut out from the gel. The digestion fragments were purified and pooled. The randomly digested DNA fragments of PAO1 and PA14 were cloned in pGM957 linearized with *Sma*I, which makes a single cut between P_{tac} and the $(SG_4)_5$-*trxA*-TIP2 chimeric ORF. The ligation was used to transform C-5907. The library was obtained by extracting plasmid DNA from the pool of clones grown in the presence of spectinomycin.

Library sequencing and data analysis

P. aeruginosa inserts cloned in the plasmid library described above were amplified by PCR with oligonucleotides 2739-2740. 10 ng of the library DNA were used in the amplification reaction. The pools of amplicons were purified by using Agencourt AMPure XP (Beckmann Coulter) in order to remove primer dimers and fragment shorter that 50 bp; afterwards, the 454 sequencing library was prepared following the Method Manual for Rapid library preparation, GS FLX Titanium (Roche Applied Science, Mannheim, Germany). 454 adaptors with MID indexes were ligated at the extremities of the library fragments. The sequencing library was analysed with Agilent Bioanalizer High Sensitivity assay and quantified with a NanoDrop fluorimeter by using PicoGreen (Invitrogen, Life Technologies). The library was sequenced in replicate in two lanes corresponding to 1/8 of a GS FLX Titanium Pico Titer Plate (PTP). Genome sequences and annotation files were retrieved

from Pseudomonas Genome Database (http://www.pseudomonas.com/). Sequencing reads were aligned to *P.aeruginosa* PAO1 and PA14 references genomes using Newbler (Roche, 454). The overlap with annotated genes and the coverage were assessed using BEDTools [40]. TSS of a gene was considered represented in the library if at least one consensus resulting from the mapping covered the start of that gene. "Core genes" were defined as genes with ≥90% identity in aminoacid sequence between PAO1 and PA14 strains using BLAST [41], otherwise they were classified as "strain-specific" genes.

β- galactosidase assays

E. coli cultures were grown at 28° in LD supplemented with ampicillin up to OD_{600}=0.5 and induced with 0.1% arabinose. The cultures were split in two and the sub-cultures were incubated at 28° and 42°. After 30 min, samples were taken to measure OD_{600} and β-galattosidase activity. BgaB activity was measured as described by Miller [42] on permeabilized cells with the modification that the assay was performed at 55 °C. The assay on *P. aeruginosa* was performed by growing the cultures at 25° in 40 ml of LD supplemented with carbenicillin (Carb) up to OD_{600}=0.5. As PAO1 cultures formed macroscopic aggregates in these conditions, the cells were collected by centrifugation and carefully resuspended in 1 mL of LD to eliminate visible aggregates. The cells were then inoculated in 40 ml of LD supplemented with Carb and 0.1% arabinose at 25°. 20 ml were immediately withdrawn and shifted at 37°. After 45 min incubation, 15 ml of each sub-culture were centrifuged and resuspended in 0.3 ml of TEDP (0.1M Tris-HCl pH 8; 1mM EDTA; 0.1M DTT; 0.1M PMSF). Crude extracts were obtained by sonication and the protein concentration was evaluated by Bradford assay [43]. β-galattosidase activity at 55 °C was assayed by mixing 0.2 ml of crude extract with 0.8 ml of Z buffer [42] and 0.2 ml of ONPG 4 g/l in Z buffer. The reaction was stopped by adding 0.5 ml of 1 M Na_2CO_3.

Northern Blotting

Procedures for RNA extraction, Northern blot analysis and *in vitro* 5'-end labeling of oligonucleotides with T4 polynucleotide kinase and [γ^{32P}] ATP were previously described [44, 45]. Oligonucleotide probes used for Northern blotting analysis were 2776 (specific for the first 23 bases of transcript from P_{tac} in pGZ119 derivatives); 2692 (TIP2); 1538 (*cat* ORF in pGZ119 and derivatives); 2865 (*bgaB*); 2809 (EGFP); 2929 (*recA*). DNA for *in vitro* transcription of *ptxS* or

dsbA riboprobes was obtained by PCR on PAO1 genomic DNA with oligonucleotides 2852-2976 and 2807-2977, respectively. Images and densitometric analysis of Northern blots were obtained by phosphorimaging using ImageQuant software (Molecular Dynamics).

Western Blotting

P. aeruginosa and *E. coli* cultures were grown in LD at either 25° or 28°, respectively, up to OD_{600} =0.5. The cultures were then induced with 0.1% arabinose and immediately split in two; the sub-cultures were then incubated 30 min at different temperatures (25°-37° for *P. aeruginosa* and 28°-42° for *E. coli*). OD_{600} was measured and 10 ml samples were collected and centrifuged 10 min at 4000 rpm. The pellet was resuspended in 100 µl of Cracking Buffer (100mM Tris-HCl pH 6.8; 200mM dithiotreitol; 4% SDS; 20% glycerol; 0,2 bromophenol blue) and boiled for 3'. After the lysis, the extracts were centrifuged at 13000 rpm for 10 min, supernatants were recovered and analyzed by western blotting. Samples corresponding to 0.0065 OD units of original cultures were run on acrylamide gel, blotted on nitrocellulose filter and hybridized with commercial antibody specific for GFP (Living Colors® A.v. Peptide Antibody, Clontech). The immunoreactive bands were revealed by Immobilon (Millipore) reagents and quantified with the ImageQuant software.

RESULTS AND DISCUSSION

The Tet-TRAP genetic tool

The Tet-Trap is a genetic tool aimed at identifying post-transcriptionally regulated genes. The system is based on Tn10 *tetA* gene regulation. In the absence of tetracycline, *tetA* transcription from *tetAp* promoter is prevented by TetR repressor binding to the operator *tetO*; tetracycline acts as an allosteric inducer of TetR. A small peptide, TIP2, is able to mimic tetracycline effect on TetR as it leads to repressor dissociation from *tetO*. TIP2 retains its properties also in end-fused chimeric polypeptides [46].

We constructed *E. coli* strains to either select for or counter-select cells expressing TIP2-tagged polypeptides (Fig. 1). In these strains, different reporter genes have been inserted downstream of *tetAp* promoter and *tetO* operator in the bacterial chromosome, between *tonB* and *yciA*. As in the reporter strains the *tetR* gene is integrated in the chromosome (between *attB* and *bioB*) and constitutively expressed from *Pcat* promoter, transcription of the reporter genes is switched off. Repression by TetR can be relieved by TIP2 and this leads to the reporter gene

expression. For positive selection of TIP2 expressing cells, we exploited as a reporter the *aadA:*GFP gene, which encodes a chimeric protein conferring spectinomycin resistance and fluorescence [31]. For negative selection purposes, we transferred the reporter cassette in a streptomycin resistant *rpsL*-31 mutant and replaced the *aadA:*GFP reporter with the *rpsL*$^+$ allele, which confers streptomycin sensitivity to the otherwise resistant strain [32]. The expression of TIP2-tagged polypeptides from a plasmid makes the two strains spectinomycin-resistant and streptomycin-sensitive, respectively (Fig. 1B).

The above *E. coli* reporter strains were preliminarily assayed for TIP2-dependent reporter gene expression by transforming them with plasmid pGM956, which carries downstream of P_{tac} an artificial gene encoding the $(SG_4)_5$-TrxA-TIP2 chimeric polypeptide (hereafter indicated as ST-TIP2). ST-TIP2 is composed by a flexible linker at the N-ter $(SG_4)_5$ followed by an *E. coli* TrxA-Tip2 fusion that has been shown to effectively inhibit TetR [46]. As negative control, all strains were also transformed with pGM957, which lacks a TIR between the promoter and the ST-TIP2 gene (Fig. 2A). As expected, only strains carrying pGM956 showed the phenotype conferred by the expression of the reporter genes upon induction of ST-TIP2 with IPTG (i.e. SpcR for *aadA:*GFP reporter and StrS for *rpsL*$^+$ (Fig. 2B).

Application of Tet-TRAP to *P. aeruginosa* RNA thermometers identification: obtaining a 5'-UTR enriched library

We applied the Tet-Trap to obtain a *P. aeruginosa* genomic library enriched for 5'-UTRs and ORFs first portions, where regions controlling translation usually map in bacteria. We cloned a mixture of random fragments of PAO1 and PA14 genomic DNA in pGM957. The cloned fragments ranged in length between 300 and 800 bp and were cloned between the P_{tac} promoter and the ST-TIP2 ORF (see Materials and Methods for experimental details). The library was used to transform the *aadA:*GFP reporter strain described above (C-5907), which allows to select for TIP2 expressing cells in the presence of spectinomycin and IPTG. Transformants were plated in permissive conditions to estimate the coverage of the genomic library and on spectinomycin and IPTG to select for translational fusions (Table 1). The library showed an high genomic coverage (more than 99% even considering PAO1 and PA14 as unrelated genomes). The 1.7% of transformants (ca. 48000 clones) were SpcR; these clones were expected to express ST-TIP2 and thus constituted the translational fusion library. We pooled all the SpcR clones, extracted plasmid DNA and amplified by PCR the inserts for 454-pyrosequencing. The 52116 reads generated by pyrosequencing, which were on average 526 bp long, were mapped on PAO1 and PA14 genomes (GenBank accession

numbers NC_002516 and NC_008463, respectively). In order to define the core and specific coding genes between PAO1 and PA14, all the protein sequences of both strains were blasted against each other; 5320 proteins showing sequence similarity equal or greater than 90% were defined as "core", 252 were identified as PAO1 specific and 572 as PA14 specific [22, 47]. We found that 3064 core coding genes, 168 PAO1-specific and 194 PA14-specific were represented in our library for a total of 3426 genes. Thus around the 60% of *P. aeruginosa* annotated coding genes (61.5% for PAO1 and 58.1% for PA14, respectively) are represented in the translational fusion library. More than 78% of these genes were sequenced with a coverage above Detection Threshold equal or greater than 3 reads per gene, thus allowing us to infer that the sequencing depth could be considered adequate to cover the whole library variability. For 2389 of the coding genes represented in the library (69.7% of the total) the cloned region encompassed the 5'-UTR and the beginning of the annotated ORF. This percentage is probably underestimated, considering the high proportion of *P. aeruginosa* genes with unknown function and no similarity to any previously reported sequences that could have been incorrectly annotated (2340 out of 5732 according to Pseudomonas Genome Database). Given the high genomic coverage of the library, we expect that missing genes may belong essentially to the following categories: i) genes physiologically poorly translated; ii) genes whose translation requires positive regulators absent in *E. coli*; iii) genes positively regulated by small molecules absent in the screening experimental conditions. Moreover, constructs encoding toxic fusion proteins would have been lost.

Application of Tet-TRAP to the identification of putative *P. aeruginosa* RNA thermometers

In the *rpsL* reporter strain, the presence of a *P. aeruginosa* RNAT in the DNA fragment cloned in frame with ST-TIP2 should result in thermosensitive expression of streptomycin resistance. Thus, no growth should be observed upon transformation of such strain (C-5920) with the translation fusion library on streptomycin at 37 °C. Conversely, by plating the transformants in the presence of the antibiotic at 28 °C, one should expect to select for clones carrying *P. aeruginosa* inserts translationally silent in these conditions and thus enrich for putative thermometers. However, upon transformation with the library, we observed an high background of StrR clones at both 28 and 37 °C, as the plating efficiency on the antibiotic was only about 250-fold lower than in the absence of selection. This was probably due to inefficient translation of a share of library fusions, as in control transformation with plasmid pGM956 we observed a more than 4.5E05-fold reduction of plating efficiency in non-permissive conditions. This implies that the amount of ST-TIP2 sufficient to make the *aadA:*GFP reporter strain (C-5907) resistant to spectinomycin may not be enough to provide

adequate induction of *rpsL*$^+$ reporter in C-5920 and lead to cell lethality on streptomycin. To distinguish clones carrying putative thermosensors, we analysed 1152 StrR transformants grown at 28 °C by replica plating on streptomycin and IPTG at 28 and 37 °C. We identified 16 clones exhibiting thermo-sensitive, IPTG-dependent streptomycin resistance (data not shown). The clones carried 12 different *P. aeruginosa* DNA inserts, as assessed by sequencing (Table 2). We tentatively classified the inserts in five groups. Four clones contained the 5'-UTR and 5'-end of ORFs in frame with ST-TIP2 (clones 1.1-to 1.4). In particular, codons 1-53 of *ptxS* (clone 1.2), 1-49 of PA5194 (1.3) and 1-34 of *dsbA* (1.4) are fused with ST-TIP2. In 1.1, the cloned ORF is annotated as an intergenic region in *Pseudomonas* database, but as it is in frame with the downstream PA1031, it could represent the actual 5'-end of such gene. Concerning the other fusions, three carried internal fragments of annotated ORFs in frame with ST-TIP2 (clones 2.1, 2.2 and 5.1, which is a chimeric construct); in other five clones, the ORFs in frame with ST-TIP2 overlapped annotated ORFs in a different frame (clones 3.1-3.3) or in antisense orientation (clones 4.1 and 4.2).

We decided to consider clones of category 1 for further analyses because they include gene regions where RNA thermometers usually map, i.e. the 5'-UTR and the first part of the coding region of the genes. However, we cannot rule out that regulatory sites may actually lie in the internal portion of ORFs. Moreover, in some cases, the assignment of our inserts to one of the above categories may have been biased by genome annotation errors. For instance, insert 4.1 encompasses codons 13-206 of a putative 349 aa long ORF, which is not annotated in pseudomonas.com database.

We estimated the efficiency of plating (eop) at 37°C *vs*. 28°C of cultures carrying the constructs 1.1-1.4 (Fig. 3). *ptxS* fusion showed the strongest effect, with a $\geq 10^5$ eop reduction at 37°C on plates supplemented with streptomycin and IPTG; for PA1031 and *dsbA*, the eop was reduced around hundredfold. Finally, for PA5194, the eop was around tenfold reduced and the colonies were very small at 37 °C. All strains showed comparable eop at the two temperatures in the absence of either streptomycin, indicating that the constructs do confer thermosensitive streptomycin resistance and not a generic thermosensitive phenotype, or IPTG. This suggests that the increase in ST-TIP2 expression was not due to temperature-dependent activation of promoters in the cloned regions, as it depends on P_{tac} induction. This was confirmed by Northern blotting analysis of the transcription pattern of the four constructs at 28 and 37 °C with a TIP2-specific probe. In all cases we did not observe any relevant difference in the transcription pattern at the two temperatures (Fig. 4). The filter was also hybridized with an oligonucleotide specific for the 5'-end of the transcript from P_{tac}, to detect short signals suggestive of transcription attenuation. Also in this case, temperature did not seemed to affect the transcription pattern. Only complete transcripts (from

P_{tac} to TIP2) were detected for constructs 1.1-1.3. For 1.4 (*dsbA*), a short signal was also present in comparable amounts at both temperatures, suggesting that a share of the transcription from P_{tac} may actually prematurely terminate within the cloned region. Visual sequence inspection of the intergenic region between *cc4* and the downstream *dsbA* genes, which is cloned in 1.4 construct, identified a palindromic region encompassing a 3'-terminal stretch of six Ts (nucleotides 6181713-6181685 in PAO1) that may reasonably correspond to the intrinsic terminator of *cc4* gene.

On the whole, the above analyses confirmed that the expression of 1.1-1.4 constructs was post-transcriptionally activated by the temperature upshift.

dsbA and _ptxS_ RNA thermometers validation in _E. coli_ with conventional reporter genes

To validate the results of the Tet-TRAP and rule out the possibility that temperature –dependent expression of 1.1-1.4 constructs was an artifact of the ST-TIP reporter system, we analysed translational fusions of two of the putative thermosensors, i.e. those identified in constructs 1.2 (*ptxS*) and 1.4 (*dsbA*), with conventional reporter genes. In particular, we exploited the *bgaB* gene, encoding a thermostable β -galactosidase [48] for *ptxS* fusions and monitored the activity of the enzyme at different temperatures. For *dsbA* validation, we applied a different strategy because DsbA has a periplasmic localization signal at the N-ter, which triggers export and thus inactivation of β-gal. We cloned *dsbA* fragments in frame with the EGFP as a reporter and monitored the expression of the fusions by Western blotting with antibodies specific for the fluorescent protein. For both *dsbA* and *ptxS* two constructs were set, i.e. a long one, carrying the whole *P. aeruginosa* regions originally found in the 1.4 and 1.2 ST-TIP2 plasmids, respectively, and a shorter one, with the same 3'-end of 1.4 and 1.2 inserts but starting with the first nucleotide transcribed from *dsbA* or *ptxS* promoters in *P. aeruginosa* [49] (FIG. 5A). Control fusions of the 5'-UTR and first 9 codons of the *E. coli recA* gene with *bgaB* and EGFP were also set. The chimeric genes were cloned in the *E. coli-P. aeruginosa* shuttle vector pGM931 and were all transcribed from the *araBp* vector promoter. The expression of the reporter genes was monitored in *E. coli* DH10B cultures at 28 and 42 °C upon transcription induction with arabinose (Table 3 and Fig. 5B).

β-gal expression in the presence of *ptxS-bgaB* fusions resulted to be strongly temperature-dependent, as cultures carrying either the long (pGM980) or the short (pGM981) constructs showed more than twentyfold or tenfold increments in the enzymatic activity at 42 °C, respectively. Conversely, BgaB activity by the control *recA* plasmid (pGM989) increased only twice. It should be noted that β-gal expression by the long pGM980 construct was much lower at both temperatures than by pGM981 (Table 3).

Similar results were obtained with DsbA-EGFP fusions. In fact, the fusion protein expressed by the long construct (pGM964) was not detectable at any temperature. On the contrary, DsbA-EGFP expressed by pGM965, which was below detection limit at 28 °C, gave a sharp signal at 42 °C. No signal was detected at 28 °C even upon over-exposition of the filter (data not shown). Thus, the expression of DsbA-EGFP was clearly thermo-dependent if compared with RecA-EGFP, which increased only around twice (Fig. 5B).

Reduced reporter expression by the long pGM980 and pGM964 fusions could be due to decreased efficiency of translation and/or stability of the mRNA transcribed from *araBp*, which bear at the 5'-end sequences that are absent in the actual *P. aeruginosa ptxS* and *dsbA* mRNAs. For *dsbA*, the putative *cc4* intrinsic terminator located between *araBp* and *dsbA* 5'-UTR may contribute to keep low the construct expression.

dsbA and *ptxS* RNATs validation in *P. aeruginosa*

The expression of *dsbA*-eGFP by plasmids pGM965 and of *ptxS-bgaB* fusion by pGM981 was analysed in PAO1 strain at 25 and 37 °C to mimic the temperature upshift due to mammal host infection.

ptxS-bgaB expression by pGM981 was found to be thermoregulated in PAO1, as it showed a more than fivefold increment at 37 °C, whereas *recA-bgaB* expression by pGM989 control plasmid was not significantly affected by the temperature upshift (Table 3). As the amount of *ptxS-bgaB* transcript did not increase at 37 °C (Fig. 6), these data indicated that *ptxS* translation is modulated in response to growth temperature.

On the other hand, when we analysed by western blotting *dsbA*-eGFP fusion expression by pGM965 in PAO1, we could not detect any signal at either temperature (Fig. 5B). Signals were absent also at 42 °C (data not shown), a temperature that allows the fusion protein expression in *E. coli* (Fig. 5B). Consistently, no signal corresponding to the *dsbA*-EGFP transcript was detected by Northern analysis (Fig. 6). Conversely, a DSBA$^{\Delta 21}$-EGFP shorter variant was expressed at comparable levels at 25 and 37 °C by plasmid pGM966 in PAO1 (Fig. 5B). This construct contains the first 36 pb of *dsbA* ORF fused with the EGFP; thus, respect to pGM965, has a 63 bp long deletion of the ORF (Fig. 5A). Notably, the deletion increased the expression of the reporter construct at both 28 and 42 °C and substantially abolished thermoregulation also in *E. coli* (around twofold increase at high temperature both for DSBA-EGFP and the control fusion RECA-EGFP; Fig. 5B). Thus, the deleted region seems to host a negative regulator of *dsbA* expression that behaves as a temperature-dependent regulatory element in *E. coli*.

Conclusions and future perspectives

In this work we have identified an RNAT regulating the expression of *ptxS* gene of *P. aeruginosa*. To our knowledge, this represents the first example of such strategy of regulation in this bacterium.

The possible physiological role of *ptxS* thermoregulation is not intuitive. The gene encodes a repressor of the LacI family that regulates transcription of its own gene and of gluconate transport and degradation operons [50, 51]. Its specific effector is 2-ketogluconate, whose binding causes the dissociation of PtxS from DNA. Recently, it has been shown that PtxS interact with nanomolar affinity with PtxR, whose gene is located adjacent and transcribed divergently from *ptxS*, and that 2-ketogluconate abolishes the interaction [52]. PtxR is a positive regulator of transcription of *toxA* gene that encodes the most toxic virulence factor of *P. aeruginosa* [53]. It has been proposed that in the absence of 2-ketogluconate, PtxS acts as a negative regulator of *toxA* by preventing the interaction of PtxR with the RNA polymerase [52]. Thus, PtxS seems to link carbon metabolism with virulence, but the role in infection of this process remains to be established. The comparison of transcriptome of *ptxS*$^{+}$ and *ptxS* mutants at 28 *vs*. 37 °C and of the virulence of such strains in the insect infection model *Galleria mellonella* at the two temperatures [54] will help in clarifying the role of thermoregulation of this gene.

No structure similarity with known RNATs has been found by mfold analysis of *ptxS* mRNA (data not shown) [55]. The identification of sequences/secondary structures involved in thermoregulation will require both *in vivo* and *in vitro* analyses, such as testing in *bgaB* reporter system the effect of nested deletions and point mutations within the first 200 nt of *ptxS* mRNA, where the RNAT is located, and structural probing experiments. Similar analyses will be performed also for *dsbA*, for which we have identified a post-transcriptional regulatory element that behaves as an RNAT in *E. coli*. As the expression of the reporter construct carrying the putative *dsbA* RNAT is almost undetectable in *P. aeruginosa,* no conclusions about thermoregulation in this bacterium can be drawn at the moment. The redox activity of the protein is required for the expression of multiple virulence factors and for the intracellular survival of *P. aeruginosa* during infection of HeLa cells [21]. Given its role in pathogenesis, it could be thus a good target for thermodependent up-regulation. DsbA expression has been documented in *P. aeruginosa* at both 30 and 37 °C [56, 57], but to our knowledge its expression levels at different temperatures have never compared so far. Northern analysis of *dsbA* transcript expressed PAO1 and PA14 showed a smeared signal with comparable intensity at both temperatures (data not shown). Irrespective on a role in thermoregulation, our data suggest that a negative *in cis* regulator of *dsbA* mRNA translation and/or stability is located in the first 180 nt of the transcript and involves sequences within *dsbA* ORF. Interestingly, DsbA is

positively regulated by two small RNAs, RsmY and RsmZ, which act by sequestering a negative post-transcriptional regulator, the RNA binding protein RsmA [58-60]. It will be interesting to assess whether the regulatory mRNA region that we have identified is involved in this regulation circuit.

The Tet-Trap seems to be a reliable method to the identification of post-transcriptionally regulated genes, at least for bacteria that follows a translation initiation pathway similar to *E. coli* one, such as other Proteobacteria. It would be interesting to apply this analysis to other bacteria, such as *E. coli* or *Salmonella* pathogenic strains. Potentially, this approach could also be applied to the identification of genes regulated by riboswitches, provided that the ligand molecule is known and can be supplied to the cells. The application of Tet-Trap to the fishing of sRNA target genes is the object of ongoing research in our lab.

ACKNOWLEDGMENTS

We thank Frank Narberhaus and Stefanie Krajewski (Ruhr-Universität Bochum – Germany) for sending pBAD2_bgaB plasmid and for helpful discussions.

REFERENCES

1. Guisbert, E., Yura, T., Rhodius, V.A. & Gross, C.A. Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev* **72**, 545-54 (2008).

2. Phadtare, S., Inouye, M. & Severinov, K. The mechanism of nucleic acid melting by a CspA family protein. *J.Mol.Biol.* **337**, 147-155 (2004).

3. Konkel, M.E. & Tilly, K. Temperature-regulated expression of bacterial virulence genes. *Microbes Infect* **2**, 157-66 (2000).

4. Johansson, J. et al. An RNA thermosensor controls expression of virulence genes in Listeria monocytogenes. *Cell* **110**, 551-61 (2002).

5. Klinkert, B. & Narberhaus, F. Microbial thermosensors. *Cell Mol Life Sci* **66**, 2661-76 (2009).

6. Waldminghaus, T., Fippinger, A., Alfsmann, J. & Narberhaus, F. RNA thermometers are common in alpha- and gamma-proteobacteria. *Biol Chem* **386**, 1279-86 (2005).

7. Kortmann, J. & Narberhaus, F. Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Microbiol* **10**, 255-65 (2012).

8. Morita, M., Kanemori, M., Yanagi, H. & Yura, T. Heat-induced synthesis of sigma32 in *Escherichia coli*: structural and functional dissection of rpoH mRNA secondary structure. *J Bacteriol* **181**, 401-10 (1999).

9. Morita, M.T. et al. Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. *Genes Dev* **13**, 655-65 (1999).

10. Nocker, A. et al. A mRNA-based thermosensor controls expression of rhizobial heat shock genes. *Nucleic Acids Res* **29**, 4800-7 (2001).

11. Altuvia, S., Kornitzer, D., Teff, D. & Oppenheim, A.B. Alternative mRNA structures of the cIII gene of bacteriophage lambda determine the rate of its translation initiation. *J Mol Biol* **210**, 265-80 (1989).

12. Giuliodori, A.M. et al. The *cspA* mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. *Mol Cell* **37**, 21-33 (2010).

13. Waldminghaus, T., Heidrich, N., Brantl, S. & Narberhaus, F. FourU: a novel type of RNA thermometer in Salmonella. *Mol Microbiol* **65**, 413-24 (2007).

14. Chowdhury, S., Maris, C., Allain, F.H. & Narberhaus, F. Molecular basis for temperature sensing by an RNA thermometer. *EMBO J* **25**, 2487-97 (2006).

15. Loh, E. et al. A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. *Cell* **139**, 770-9 (2009).

16. Kortmann, J., Sczodrok, S., Rinnenthal, J., Schwalbe, H. & Narberhaus, F. Translation on demand by a simple RNA-based thermosensor. *Nucleic Acids Res* **39**, 2855-68 (2011).

17. Stover, C.K. et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**, 959-64 (2000).

18. Pier, G. Application of vaccine technology to prevention of *Pseudomonas aeruginosa* infections. *Expert Rev Vaccines* **4**, 645-56 (2005).

19. Wurtzel, O. et al. The single-nucleotide resolution transcriptome of *Pseudomonas aeruginosa* grown in body temperature. *PLoS Pathog* **8**, e1002945 (2012).

20. Colmer, J.A. & Hamood, A.N. Characterization of *ptxS*, a *Pseudomonas aeruginosa* gene which interferes with the effect of the exotoxin A positive regulatory gene, *ptxR*. *Mol Gen Genet* **258**, 250-9 (1998).

21. Ha, U.H., Wang, Y. & Jin, S. DsbA of *Pseudomonas aeruginosa* is essential for multiple virulence factors. *Infect Immun* **71**, 1590-5 (2003).

22. Lee, D.G. et al. Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biol* **7**, R90 (2006).

23. Rahme, L.G. et al. Common virulence factors for bacterial pathogenicity in plants and animals. *Science* **268**, 1899-902 (1995).

24. Bertrand, K.P., Postle, K., Wray, L.V., Jr. & Reznikoff, W.S. Overlapping divergent promoters control expression of Tn10 tetracycline resistance. *Gene* **23**, 149-56 (1983).

25. Sasaki, I. & Bertani, G. Growth abnormalities in Hfr derivatives of *Escherichia coli* strain C. *J.Gen.Microbiol.* **40**, 365-376 (1965).

26. Carzaniga, T. et al. Autogenous regulation of *Escherichia coli* polynucleotide phosphorylase expression revisited. *J.Bacteriol.* **191**, 1738-1748 (2009).

27. Yanisch-Perron, C., Vieira, J. & Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-19 (1985).

28. Sukhodolets, M.V. & Garges, S. Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42**, 8022-8034 (2003).

29. Regonesi, M.E. et al. A mutation in polynucleotide phosphorylase from *Escherichia coli* impairing RNA binding and degradosome stability. *Nucleic Acids Res.* **32**, 1006-1017 (2004).

30. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc.Natl.Acad.Sci.U.S.A* **97**, 6640-6645 (2000).

31. Rizzi, A. et al. Strategy for in situ detection of natural transformation-based horizontal gene transfer events. *Appl Environ Microbiol* **74**, 1250-4 (2008).

32. Lederberg, J. Streptomycin resistance; a genetically recessive mutation. *J Bacteriol* **61**, 549-50 (1951).

33. Georgi, C., Buerger, J., Hillen, W. & Berens, C. Promoter strength driving TetR determines the regulatory properties of Tet-controlled expression systems. *PLoS One* **7**, e41620 (2012).

34. Lessl, M. et al. Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region in trans. *J.Bacteriol.* **174**, 2493-2500 (1992).

35. Briani, F., Ghisotti, D. & Dehò, G. Antisense RNA-dependent transcription termination sites that modulate lysogenic development of satellite phage P4. *Mol.Microbiol.* **36**, 1124-1134 (2000).

36. Carzaniga, T., Antoniani, D., Dehò, G., Briani, F. & Landini, P. The RNA processing enzyme polynucleotide phosphorylase negatively controls biofilm formation by repressing poly-N-acetylglucosamine (PNAG) production in *Escherichia coli* C. *BMC.Microbiol.* **12**, 270 (2012).

37. Qiu, D., Damron, F.H., Mima, T., Schweizer, H.P. & Yu, H.D. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. *Appl.Environ.Microbiol.* **74**, 7422-7426 (2008).

38. Klinkert, B. et al. Thermogenetic tools to monitor temperature-dependent gene expression in bacteria. *J Biotechnol* **160**, 55-63 (2012).

39. Goldberg, J.B. & Ohman, D.E. Cloning and expression in *Pseudomonas aeruginosa* of a gene involved in the production of alginate. *J Bacteriol* **158**, 1115-21 (1984).

40. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-2 (2010).

41. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-10 (1990).

42. Miller, J.H. Experiments in molecular genetics (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1972).

43. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).

44. Dehò, G., Zangrossi, S., Sabbattini, P., Sironi, G. & Ghisotti, D. Bacteriophage P4 immunity controlled by small RNAs *via* transcription termination. *Mol Microbiol* **6**, 3415-25 (1992).

45. Briani, F. et al. Genetic analysis of polynucleotide phosphorylase structure and functions. *Biochimie* **89**, 145-157 (2007).

46. Goeke, D. et al. Short peptides act as inducers, anti-inducers and corepressors of Tet repressor. *J Mol Biol* **416**, 33-45 (2012).

47. Silby, M.W., Winstanley, C., Godfrey, S.A., Levy, S.B. & Jackson, R.W. Pseudomonas genomes: diverse and adaptable. *FEMS Microbiol Rev* **35**, 652-80 (2011).

48. Hirata, H., Negoro, S. & Okada, H. Molecular basis of isozyme formation of beta-galactosidases in *Bacillus stearothermophilus*: isolation of two beta-galactosidase genes, bgaA and bgaB. *J Bacteriol* **160**, 9-14 (1984).

49. Dotsch, A. et al. The *Pseudomonas aeruginosa* transcriptome in planktonic cultures and static biofilms using RNA sequencing. *PLoS One* **7**, e31092 (2012).

50. Swanson, B.L., Colmer, J.A. & Hamood, A.N. The *Pseudomonas aeruginosa* exotoxin A regulatory gene, ptxS: evidence for negative autoregulation. *J Bacteriol* **181**, 4890-5 (1999).

51. Swanson, B.L. & Hamood, A.N. Autoregulation of the *Pseudomonas aeruginosa* protein PtxS occurs through a specific operator site within the ptxS upstream region. *J Bacteriol* **182**, 4366-71 (2000).

52. Daddaoua, A. et al. Genes for carbon metabolism and the ToxA virulence factor in *Pseudomonas aeruginosa* are regulated through molecular interactions of PtxR and PtxS. *PLoS One* **7**, e39390 (2012).

53. Hamood, A.N., Colmer, J.A., Ochsner, U.A. & Vasil, M.L. Isolation and characterization of a *Pseudomonas aeruginosa* gene, ptxR, which positively regulates exotoxin A production. *Mol Microbiol* **21**, 97-110 (1996).

54. Jander, G., Rahme, L.G. & Ausubel, F.M. Positive correlation between virulence of *Pseudomonas aeruginosa* mutants in mice and insects. *J.Bacteriol.* **182**, 3843-3845 (2000).

55. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406-15 (2003).

56. Imperi, F. et al. Analysis of the periplasmic proteome of *Pseudomonas aeruginosa*, a metabolically versatile opportunistic pathogen. *Proteomics* **9**, 1901-15 (2009).

57. Urban, A., Leipelt, M., Eggert, T. & Jaeger, K.E. DsbA and DsbC affect extracellular enzyme formation in *Pseudomonas aeruginosa*. *J Bacteriol* **183**, 587-96 (2001).

58. Kay, E. et al. Two GacA-dependent small RNAs modulate the quorum-sensing response in *Pseudomonas aeruginosa*. *J.Bacteriol.* **188**, 6026-6033 (2006).

59. Burrowes, E., Abbas, A., O'Neill, A., Adams, C. & O'Gara, F. Characterisation of the regulatory RNA RsmB from *Pseudomonas aeruginosa* PAO1. *Res Microbiol* **156**, 7-16 (2005).

60. Heurlier, K. et al. Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. *J.Bacteriol.* **186**, 2936-2945 (2004).

TABLES

Table 1. *P. aeruginosa* genomic library features

Library size[a]	2.8E06
Genomic coverage[b]	>99%
Translational fusions[c]	4.8E04

[a] The total number of clones constituting the genomic library, as estimated by plating an aliquot of the transformants on chloramphenicol. 2.0E03 CamR clones were obtained in the control ligation (with vector only)
[c]Clones grown on spectinomycin and IPTG

Table 2. *P. aeruginosa* regions carrying putative thermosensors

Insert	n.[a]	PAO1 co-ordinates[b]	Genome annotation[d]	ORF in the construct in frame with ST-TIP2[e]
1.1	4	1117570-1117919	PA1030.1 (+; 1)-IR	264 bp at the 5'-end of an ORF not annotated in Pseudomonas.com database; the ORF is in frame with the downstream PA1031 (+; 4) ORF
1.2	1	2487532-2488014	IR- *ptxS* (+; 1)	codons 1-53of *ptxS*
1.3	1	5846939-5847277	*yrfI* (-; 2) -IR-PA5194 (+; 4)	codons 1-49 of PA5194
1.4	1	6181479-6181766[c]	*cc4* (-; 2)-IR-*dsbA* (-; 2)	codons 1-34 of *dsbA*
2.1	1	4659253-4659565	PA4163 (+; 4)	303 bp long ORF corresponding to codons 272-372 of PA4163
2.2	1	5614540-5614840[c]	*msbA* (-; 2)	300 bp long ORF corresponding to codons 236-335 of *msbA*
3.1	1	3252480-3252988[c]	PA2897 (-; 3)	75 bp long ORF corresponding to an internal fragment of PA2897 in a different frame
3.2	1	4313365-4313633 [c]	PA3852 (-; 4)	195 bp long ORF corresponding to an internal fragment of PA3852 in a different frame
3.3	2	5124494-5124786[c]	PA4576 (-; 3)	213 bp long ORF corresponding to an internal fragment of PA4576 in a different frame
4.1	1	1746834-1747420	PA1604 (-; 4)	codons 13-206 of a not annotated 1047 bp long ORF overlapping PA1604 in antisense orientation
4.2	1	3799519-3799817 [c]	*nosD* (+; 2)	codons 117-215 of a not annotated 1059 bp long ORF overlapping *nosD* in antisense orientation
5.1	1	4455043-4455361+ 2996575-2996915	*ladS* (-; 1) + *nuoN* (+; 2)	656 bp chimeric ORF; the 3'-end corresponds to codons 105-199 of *nuoN*

[a] Number of sequenced clones carrying the insert

[b] Co-ordinates refer to GenBank accession number NC_002516

[c] Region cloned in antisense orientation

[d] Intergenic regions (IR) and genes (partially) overlapping *Pseudomonas* regions cloned upstream of ST-TIP2 are indicated in the same order in which they are in the corresponding construct (from *Ptac* to ST-TIP2). For the genes, the strand and the confidence rating assigned to the predicted gene function are reported in brackets. 1, genes of known function in *P. aeruginosa*; 2, similarity with

well-characterized genes from other bacteria; 3, presence of known functional motifs; 4, unknown function [17] (www.pseudomonas.com).

[e] The length of the longest predicted ORF in frame with ST-TIP2 is indicated. Whenever an annotated ORF is in frame with ST-TIP2, the region from the initiation codon is reported.

Table 3. *ptxS* RNAT validation

	E. coli[a]			PAO1[a]		
	28°	42°	IF[b]	25°	37°	IF[b]
pGM980	3	69	23.3±2.4	nt	nt	na
pGM981	61	827	12.6±1.2	145	787	5.2±0.2
pGM989	855	2115	2.6±0.2	3447	4853	1.4±0.4

[a] *bgaB* activity (Miller units) monitored as described in Materials and Methods. The results of typical experiments are reported.

[b] Induction Factor of *bgaB* activity at high *vs*. low temperature. The values are the average of two independent determinations.

FIGURE LEGENDS

Fig. 1. *E. coli* reporter strains of Tet-Trap system. A. Schematic representation of the constructs integrated into the *E. coli* chromosome in Tet-Trap reporter strains. The reporter cassette is integrated between nucleotides 1309870 -1309872; *tetR* cassette is integrated between 806594-808521. Details about strain construction and elements composing the cassettes are reported in Materials and Methods. Construct elements are reported on an arbitrary scale. Empty squares, open reading frames; bent arrows, promoters; triangles, intrinsic bi-directional terminators[35]. B. Reporter gene expression regulation in the Tet-Trap system. The reporter strains carry genes conferring either spectinomycin resistance (*aadA:*GFP gene; upper part, strain C-5907) or streptomycin sensitivity (*rpsL* gene; lower part, strain C-5920) downstream of *tetAp* promoter. In both strains, *tetR* gene is constitutively expressed from P_{cat} and transcription from *tetAp* is repressed. In C-5920, the endogenous *rpsL* allele carries the *rpsL-31* recessive mutation, conferring streptomycin-resistance. Repression by TetR is relieved by TIP2 expression that inhibits DNA binding by TetR and triggers reporter genes transcription. Grey ovals, TetR; dark grey bar, TIP2 peptide; thick arrows, mRNAs. Other symbols are as in A.

Fig. 2. Probing the Tet-Trap system with TIP2 expressing plasmids. A. Map of plasmid encoding ST-TIP2. Details about plasmid construction and coordinates of the cloned regions are reported in Materials and Methods. Dotted lines, vector sequence; bent arrows, promoter. The constructs carry the IPTG inducible promoter P_{tac}. The empty box represent the ST-TIP2 ORF; the grey box is the TIR region. The asterisk indicates the *Sma*I cloning site used for library construction. B. Plating efficiency (eop) of *aadA*:GFP (C-5907; upper panels) and *rpsL* (C-5920; lower panels) reporter strains carrying either pGM956 or pGM957. Serial 10-fold dilutions of overnight cultures carrying the plasmid indicated were replicated on LD+ chloramphenicol plates with or without spectinomycin (Spc) and IPTG and incubated for 16 h at 37 °C.

Fig. 3. Plating efficiency (eop) at 28 *vs.* 37 °C of clones carrying putative RNATs. Serial 10-fold dilutions of C-5920 overnight cultures carrying putative *P. aeruginosa* RNATs upstream of ST-TIP2 were replicated on LD+ chloramphenicol plates in the presence or absence of streptomycin (Str) and IPTG and incubated for 16-20 h at the indicated temperatures. C-5920 carrying pGM956

and pGM957were used as positive and negative controls, respectively, of ST-TIP2 dependent *rpsL*$^{+}$ expression.

Fig. 4. Northern analysis of putative RNATs transcription in *E. coli.* The RNA was extracted from exponential cultures of C-5920 grown at 28 or 37 °C (as indicated below the panels). 20 µg of RNA were loaded on a denaturing 1.5% agarose gel and analyzed by Northern blotting with the following radiolabelled oligonucleotides: 2692 (TIP2-specific, upper panels), 2776 (complementary to the 5'-end of the transcripts from P_{tac}, central panels); 1835 (loading control; complementary to the vector *cat* gene, lower panels). *P. aeruginosa* region cloned upstream of ST-TIP2 in each strain is reported on top of the panels.

Fig. 5. Validation of putative RNATs. A. Map of plasmid encoding *dsbA*-GFP and *ptxS-bgaB* variants. Details about plasmid construction and coordinates of the cloned regions are reported in Materials and Methods. *P. aeruginosa* 5′-UTRs and ORFs are drawn to scale. Dotted lines, vector sequences; bent arrows, *araBp* promoter; empty boxes, *P. aeruginosa* DNA; grey boxes, reporter genes. B. Western analysis of *dsbA*-GFP and *recA*-GFP expression. *E. coli* DH10B (left) or *P. aeruginosa* PAO1 (right) cultures carrying the plasmids indicated above the lanes were grown and proteins extracted as described Materials and Methods. Proteins were separated by 10% denaturing polyacrylamide gel electrophoresis, blotted onto nitrocellulose filters and immunodecorated with antibodies specific for the GFP or, as a loading control, for L4.

Fig. 6. Northern blot analysis of *dsbA*-GFP and *ptxS-bgaB* fusions in *P. aeruginosa.* Cultures of PA01 strain carrying the plasmids indicated above the lanes were grown 25° or 28°, as indicated, up to OD_{600}=0.5. The cultures were induced with 0.1% arabinose, split and further incubated at the temperature indicated above the lanes for 45 min. Northern blotting was performed as described in Materials and Methods upon 1.5% denaturing agarose gel electrophoresis. The filter was hybridized with oligonucleotides specific for *bgaB* (left) or EGFP (right).

A


tetRp tetAp
P cat
lacZα
reporter gene
tetR
tonBt
tetO1 tetO2
tΩ
tI3

B

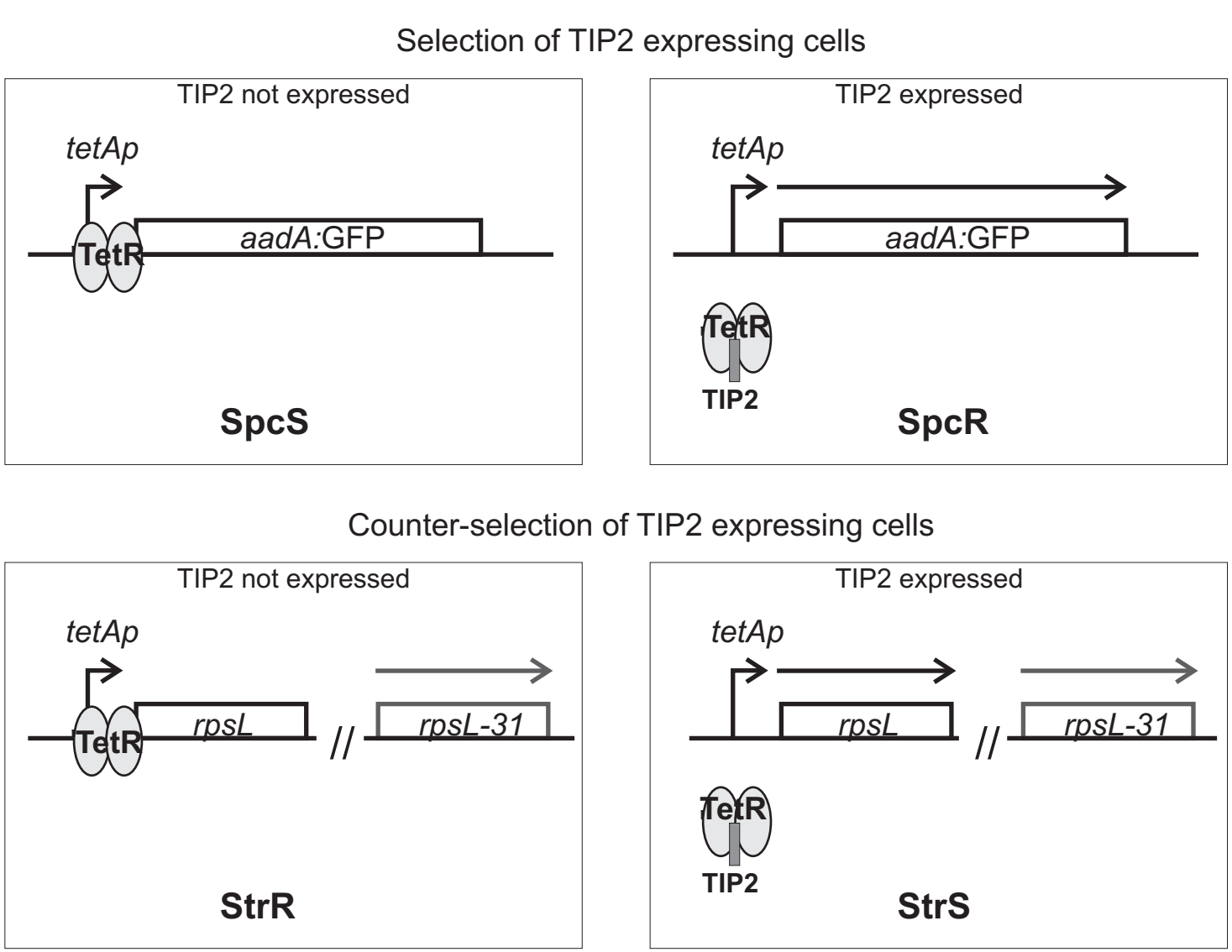

Selection of TIP2 expressing cells
TIP2 not expressed
tetAp
TetR
aadA:GFP
SpcS
TIP2 expressed
tetAp
aadA:GFP
TetR
TIP2
SpcR
Counter-selection of TIP2 expressing cells
TIP2 not expressed
tetAp
TetR
rpsL
rpsL-31
StrR
TIP2 expressed
tetAp
rpsL
rpsL-31
TetR
TIP2
StrS

Fig. 1

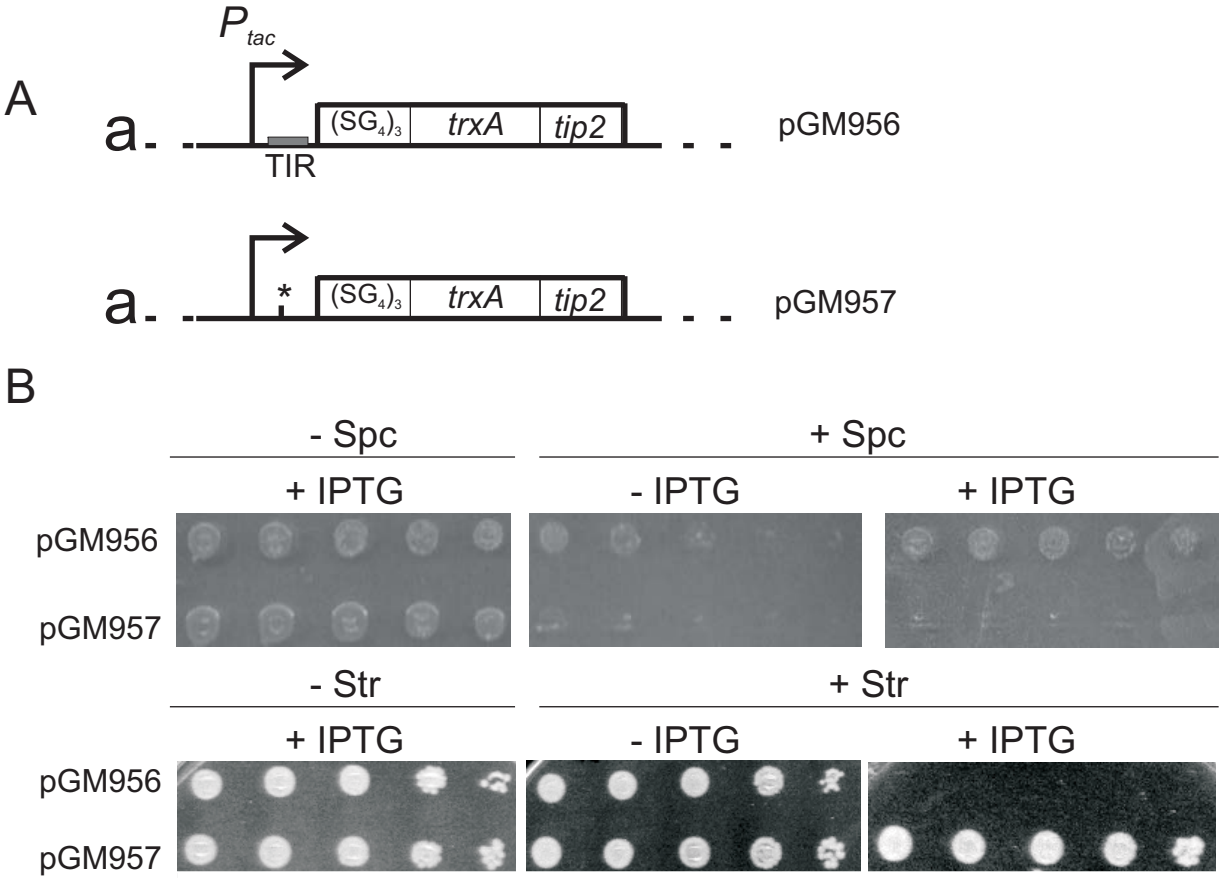

A
P_{tac}
a
$(SG_4)_3$
trxA
tip2
TIR
pGM956
a
*
$(SG_4)_3$
trxA
tip2
pGM957
B
- Spc
+ Spc
+ IPTG
- IPTG
+ IPTG
pGM956
pGM957
- Str
+ Str
+ IPTG
- IPTG
+ IPTG
pGM956
pGM957

Fig. 2

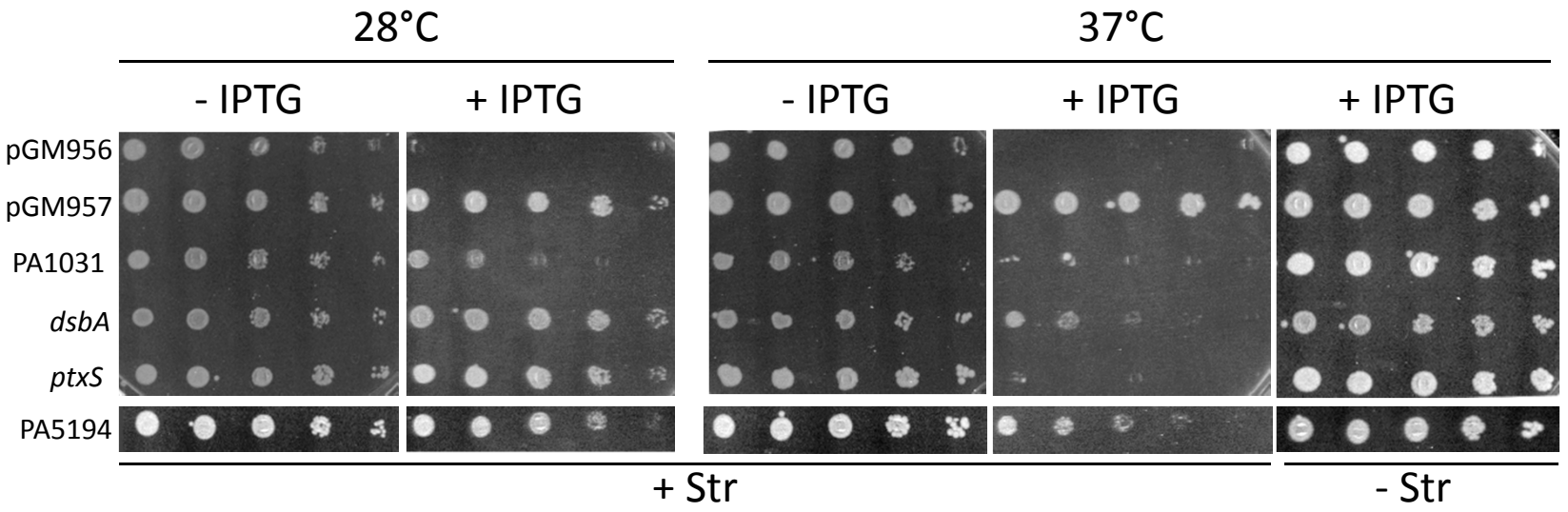

28°C
37°C
- IPTG
+ IPTG
- IPTG
+ IPTG
+ IPTG
pGM956
pGM957
PA1031
dsbA
ptxS
PA5194
+ Str
- Str

Fig. 3

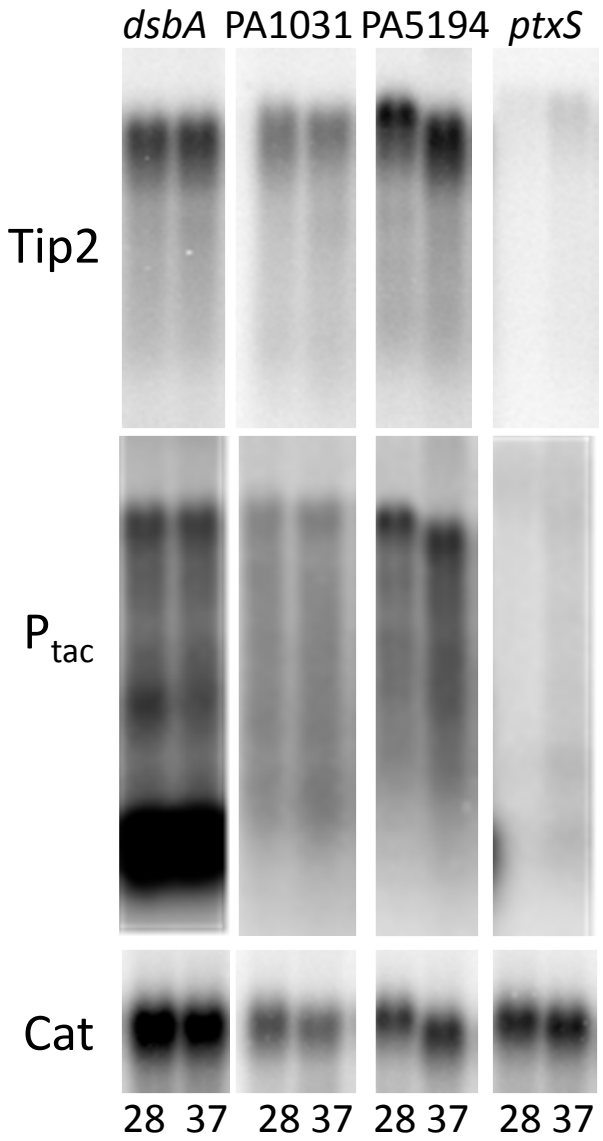


Fig. 4

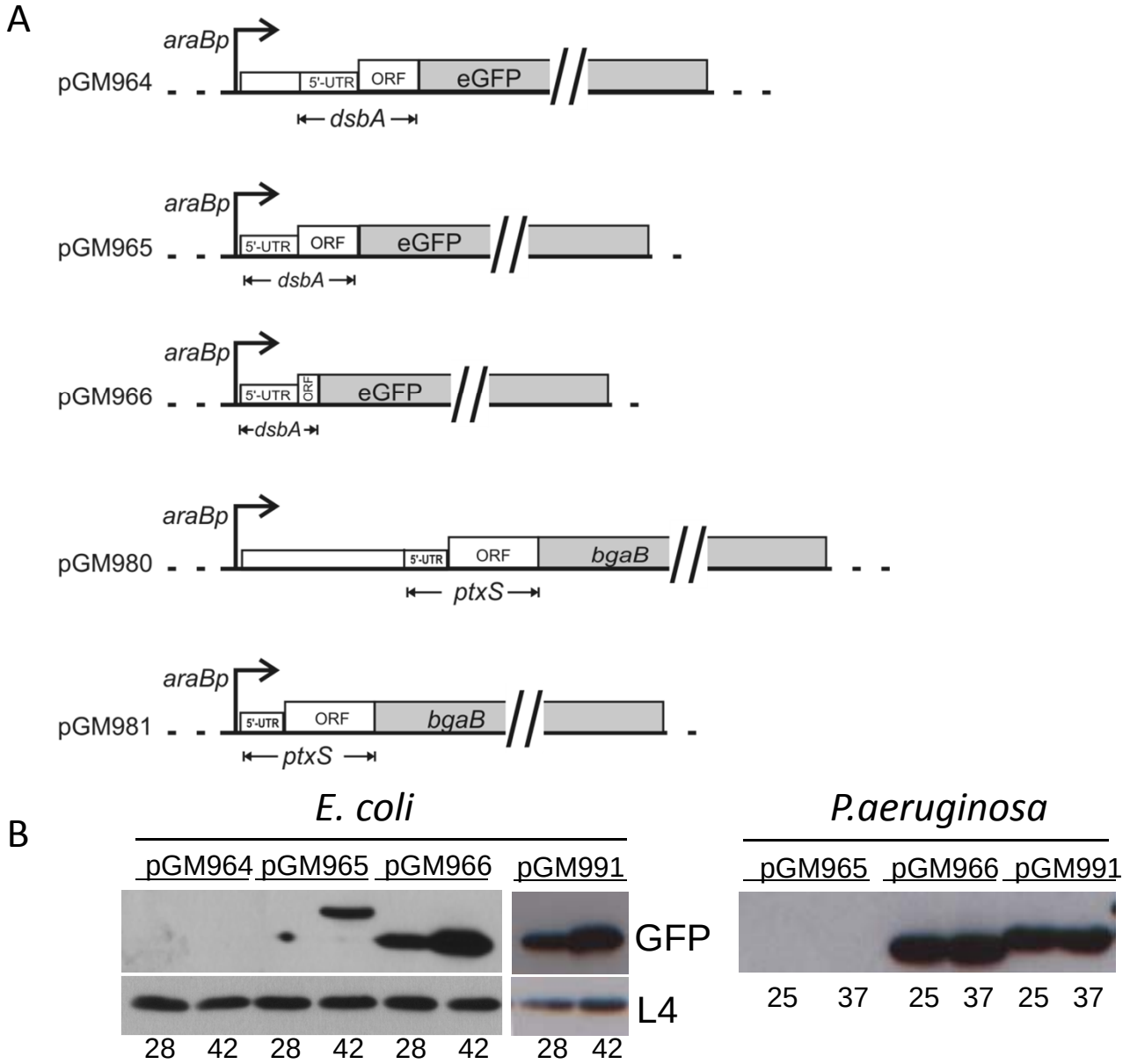

A
araBp
pGM964
5'-UTR
ORF
eGFP
dsbA
pGM965
pGM966
pGM980
bgaB
ptxS
pGM981
B
E. coli
P.aeruginosa
pGM964 pGM965 pGM966 pGM991
pGM965 pGM966 pGM991
GFP
L4
28 42 28 42 28 42 28 42
25 37 25 37 25 37

Fig. 5

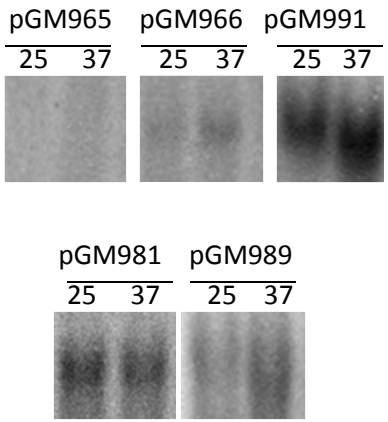

pGM965
25 37
pGM966
25 37
pGM991
25 37
pGM981
25 37
pGM989
25 37

Fig. 6

Supplementary Table S1. Bacterial strains, plasmids, phages and oligonucleotides

Strain	Relevant Genotype	Origin or reference
Pseudomonas aeruginosa		
PA01		[1]
PA14		[2]
Escherichia coli		
C-1a	*E. coli C*, prototrophic	[3]
C-5708	C-1a *rpsL-31*	laboratory collection
C-5868	C-1a Δ*pnp-751* Δ*bcsA*::*cat tn10*	[4]
C-5898	C-1a *tetRp αlacZ tetAp-kanR*	this work
C-5899	C-1a *tetRp-αlacZ tetAp-aadgfp*	this work
C-5901	C-1a *Δbio tetRp- αlacZ tetAp-aadgfp Pcat*$^{+}$*-tetR:kan*^{R}	this work
C-5907	C-1a *Δbio tetRp- αlacZ tetAp-aadgfp Pcat*$^{+}$*-tetR*	this work
C-5912	C-1a *tetRp- αlacZ tetAp-rpsL*$^{+}$*:cat*	this work
C-5916	C-5708 *tetRp-αlacZ tetAp-rpsL*$^{+}$*::cat*	this work
C-5918	C-5708 *tetRp- αlacZ tetAp-rpsL*$^{+}$	this work
C-5920	C-5708 *tetRp- αlacZ tetAp-rpsL*$^{+}$ *Pcat*-10^{CATTTA} -*tetR:kan*^{R}	this work
BW25113	*E. coli* K-12	[5]
KG264	BW25113 *Pcat*$^{+}$*-tetR:kan*^{R}	this work
KG265	BW25113 *Pcat*-10^{CATTTA} *-tetR:kan*^{R}	this work

Plasmids and phage	Relevant characteristics	Reference
pBAD2-bgaB	carries *Bacillus stearothermophilus bgaB* gene	
pBAD24-Δ1	pBAD24 derivative	[6]
pGM362	carries pHP45 t_{Ω} terminator	[7]
pGM742	*oriVColD; CamR*	[8]
pGM930	pBAD24-Δ1 derivative carrying pHP45 t_{Ω} terminator downstream of *araBp*	this work
pGM931	pHERD20T derivative carrying *araBp*- t_{Ω} region of pGM930	this work
pGM932	pGM742 derivative carrying the *lacZ*α- *tetRp-tetO-tetAp- kan*^{R} cassette	this work
pGM956	pGZ119HE derivative. Carries pQE31-S1 Shine-Dalgarno and ATG in frame with ST-TIP2.	this work
pGM957	pGZ119HE derivative. Carries pQE31-S1 Shine-Dalgarno and ATG out of frame with ST-TIP2.	this work
pGM963	pGM931 derivative with the insertion of EGFP	this work
pGM964	pGM963 derivative, carries *dsbA* (6318732-6318445) translationally fused to EGFP.	this work
pGM965	pGM963 derivative, *dsbA* (6318624-6318445) translationally fused to EGFP.	this work
pGM966	pGM963 derivative, carries *dsbA* (6318624-6318507) translationally fused to EGFP.	this work
pGM978	pGM931 derivative, contains *bgaB* under pBAD control	this work
pGM980	pGM978 derivative, carries *ptxS* (2487532-2488013) translationally fused to *bgaB* .	this work
pGM981	pGM978 derivative, carries *ptxS* (2487779-2488013) translationally fused to *bgaB*	this work
pGM982	pGM978 derivative, carries *ptxS* (2487779-2487879) translationally fused to *bgaB*	this work
pGM989	pGM978 derivative, carries *recA* (2334354-2334277) translationally fused to *bgaB*	this work
pGM991	pGM963 derivative, carries *recA* (2334354-2334277) translationally fused to EGFP.	this work

pGZ119HE	oriVColD; CamR ; Ptac	[9]
pHERD20T	*P. aeruginosa-E. coli* shuttle vector	[10]
pKD46	carries λ RED recombination genes	[5]
pQE31S1	AmpR, ColE1; *rpsA* under P_{tac} promoter	[11]
pUC19	AmpR, ColE1	[12]
pWH2354		[13]
pZR80-2	carries the chimeric aadA::gfp gene	[14]
P1 HTF	High transduction frequency phage P1 derivative	[15]

Oligo	Sequence[a]
1538	GCTGAACGGTCTGGTTATAG
2600	CTATCAGTGATAGAGAAAAGTGAAATGATTGAACAAGATGGATTG
2601	TAGTCTCGGTCCCCCATAAAAAAGGGACCTCTAGGGTCCCCAAGTCGGTCA TTTCGAACCCC
2602	CATTAATTCCTAATTTTTGTTGACAC
2603	TTCACTTTTCTCTATCACTGATAG
2604	GAAATTCAGTAAAAGCCTCCGACCGGAGGCTTTTGACTGGCGGGTGTCGGG GCTG
2605	GTGTCAACAAAAATTAGGAATTAATGA TGACCATGATTACGCCAAGC
2606	ATGTCGCGGTTGATCCTGAAGGAAAAC CTC
2617	AGCTTATTAAAGAGGAGAAATTAACTA TGAGAGGCATG
2618	CCTCTCATAGTTAATTTCTCCTCTTTAATA
2636	GGTTCGAAATGACCGACTTGGGGACCCTAGAGGTCCCTTTTTTATGGGGGG TGTAGGCTGGAGCTGCTT
2638	TTGATCCTGAAGGAAAAACCTCGCGCCTTACCTGTTGAGTAATAGTCTCGGT TAAAAAATGCCCTCTTGGGTTA
2683	TGATAGAGTTATTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAATGG ATCCCGAA GCGGTG
2684	GACCTCTAGGGTCCCCAAGTCGGTCATTTCGAACCCCAGAGTCCCGCTCAT GATGCCTGGAATTAATTCC
2685	AGCCTGCTTTTTTATACTAACTTGAGCGAAACGGGAAGGTAAAAAGACAAC TTCGTCTGTTTCTACTGG
2686	CCATGGGGCTTCTCCAAAACGTGTTTTTTGTTGTTAATTCGGTGTAGACTTT GTGTAGGCTGGAGCTGC
2689	ACCCGGGAGTGGTGGTGGCGGCAGCGGCGGTGGTGGATCCGGTGGCGGTG GCTC
2690	TACCGGTACCGCCGCCACCCGAACCGCCACCGCCAGAGCCACCGCCACCGG A

2691 GGCGGCGGTACCGGTAGCGATAAAATTATTCACCTGACT

2692 CCC**GAATTC**CGTTACCAATGCCACATCCAC AT

2693 CCCTGCAGCATGCAAACCCGGGAGTGGTG GT

2712 TAAAAAATGCCCTCTTGGGTTACATATGAA TATCCTCCTTAGT

2713 TCAGTGATAGAGAAAAGTGAAATGGCAAC AGTTAACCAGC

2714 AAGTCGGTCATTTCGAACCTTACTTAACGGA GAACC

2776 TAATTCTGTTTCCTGTGTGAAAT

2803 CTC**GGTACC**AGTAAAGGAGAAGAACTTTTCAC

2804 CTC**CTGCAG**CTATTTGTATAGTTCATCCATGC

2806 ACC**GGTACC**CTACATCCAGGGTCTGCACT

2807 ACC**GGTACC**ACCCGTGACCGCTCCCCC

2808 ACC**GGTACC**CTCGACGTATTCCTTGCCGG

2809 ACC**GGTACC** CATGGCCAGCATGGCGGT

2846 GCG**CCATGG** GGGAGCATATGCGAATCTTC

2847 CCCTGATCTCGACCTGCA

2850 GGG**CCATGG**CTCGCCGATTGATCGCTTTC

2851 CCC**GAATTC**ATGGTCGATGGCGCGCTC

2852 GGG**CCATGG**GGTTTCAACTCCTGGCATCC

2853 CCC**GAATTC**GCTGGGCAGTACTGAACC

2865 TCAACTTAGCATCTTCATACC

2915 TCT**CCATGG**CAACAGAACATATTGACTATCC

2916	CAC**GAATTC**TTTCTGTTTGTTTTCGTCGATAG
2928	TCT**GGTACC**CAACAGAACATATTGACTATCC
2929	CAC**GGTACC**TTTCTGTTTGTTTTCGTCGATAG
2976	*CTAATACGACTCACTATAG*GGCCTCGGCCACCTGGTT
2977	*CTAATACGACTCACTATAG*GGCCAGCATGGCGGTGAG

[a] Boldface characters, restriction sites; italics, T7 promoter

SUPPLEMENTARY REFERENCES

1. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M *et al*: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* 2000, 406(6799):959-964.
2. Rahme LG, Stevens EJ, Wolfort SF, Shao J, Tompkins RG, Ausubel FM: Common virulence factors for bacterial pathogenicity in plants and animals. *Science* 1995, 268(5219):1899-1902.
3. Sasaki I, Bertani G: Growth abnormalities in Hfr derivatives of *Escherichia coli* strain C. *JGenMicrobiol* 1965, 40(3):365-376.
4. Carzaniga T, Briani F, Zangrossi S, Merlino G, Marchi P, Dehò G: Autogenous regulation of *Escherichia coli* polynucleotide phosphorylase expression revisited. *JBacteriol* 2009, 191(6):1738-1748.
5. Datsenko KA, Wanner BL: One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *ProcNatlAcadSciUSA* 2000, 97(12):6640-6645.
6. Carzaniga T, Antoniani D, Dehò G, Briani F, Landini P: The RNA processing enzyme polynucleotide phosphorylase negatively controls biofilm formation by repressing poly-N-acetylglucosamine (PNAG) production in *Escherichia coli* C. *BMCMicrobiol* 2012, 12(1):270.
7. Briani F, Ghisotti D, Dehò G: Antisense RNA-dependent transcription termination sites that modulate lysogenic development of satellite phage P4. *MolMicrobiol* 2000, 36(5):1124-1134.
8. Regonesi ME, Briani F, Ghetta A, Zangrossi S, Ghisotti D, Tortora P, Dehò G: A mutation in polynucleotide phosphorylase from Escherichia coli impairing RNA binding and degradosome stability. *Nucleic Acids Res* 2004, 32(3):1006-1017.
9. Lessl M, Balzer D, Lurz R, Waters VL, Guiney DG, Lanka E: Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region in trans. *JBacteriol* 1992, 174(8):2493-2500.
10. Qiu D, Damron FH, Mima T, Schweizer HP, Yu HD: PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. *ApplEnvironMicrobiol* 2008, 74(23):7422-7426.
11. Sukhodolets MV, Garges S: Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* 2003, 42(26):8022-8034.
12. Yanisch-Perron C, Vieira J, Messing J: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 1985, 33(1):103-119.
13. Georgi C, Buerger J, Hillen W, Berens C: Promoter strength driving TetR determines the regulatory properties of Tet-controlled expression systems. *PloS one* 2012, 7(7):e41620.
14. Rizzi A, Pontiroli A, Brusetti L, Borin S, Sorlini C, Abruzzese A, Sacchi GA, Vogel TM, Simonet P, Bazzicalupo M *et al*: Strategy for in situ detection of natural transformation-based horizontal gene transfer events. *Applied and environmental microbiology* 2008, 74(4):1250-1254.
15. Wall JD, Harriman PD: Phage P1 mutants with altered transducing abilities for *Escherichia coli Virology* 1974, 59(2):532-544.