A computational strategy for the search of regulatory small RNAs in *Actinobacillus pleuropneumoniae*

CIRO C. ROSSI[1], JANINE T. BOSSÉ[2], YANWEN LI[2], ADAM A. WITNEY[3], KATE A. GOULD[3], PAUL R. LANGFORD[2] and DENISE M. S. BAZZOLLI[1]

[1]Laboratório de Genética Molecular de Micro-organismos, Departamento de Microbiologia, Instituto de Biotecnologia Aplicada à Agropecuária – BIOAGRO, Universidade Federal de Viçosa, Viçosa, 36570-900, Brazil

[2]Section of Paediatrics, Imperial College London, St Mary's Campus, London W2 1PG, UK

[3]Institute for Infection and Immunity, St George's, University of London, Cranmer Terrace, London SW17 0RE, UK

Running head: Computational sRNA search in *A. pleuropneumoniae*

Keywords: sRNAs; *Pasteurellaceae*; Porcine pleuropneumonia; bioinformatics; virulence

Corresponding authors: Denise M. S. Bazzolli (dbazzolli@ufv.br) and Paul R. Langford (p.langford@imperial.ac.uk)

ABSTRACT

Bacterial regulatory small RNAs (sRNAs) play important roles in gene regulation and are frequently connected to the expression of virulence factors in diverse bacteria. Only a few sRNAs have been described for *Pasteurellaceae* pathogens and no sRNA has yet been described for *Actinobacillus pleuropneumoniae*, the causative agent of porcine pleuropneumonia, responsible for considerable losses in the swine industry. To search for sRNAs in *A. pleuropneumoniae*, we developed a pipeline for the computational analysis of the bacterial genome by using four algorithms with different approaches, followed by experimental validation. The coding strand and expression of 17 out of 23 RNA candidates were confirmed by Northern blotting, RT-PCR and RNA sequencing. Among them, two are likely riboswitches, three are housekeeping regulatory RNAs, two are the widely studied GcvB and 6S sRNAs, and ten are putative novel *trans*-acting sRNAs, never before described for any bacteria. The latter group has several potential mRNAs targets, many of which are involved with virulence, stress resistance or metabolism, and connect the sRNAs in a complex gene regulatory network. The sRNAs identified are well conserved among the *Pasteurellaceae* that are evolutionarily closer to *A. pleuropneumoniae* and/or share the same host. Our results show that the combination of newly developed computational programs can be successfully utilized for the discovery of novel sRNAs and indicate an intricate system of gene regulation through sRNAs in *A. pleuropneumoniae* and in other *Pasteurellaceae*, thus providing clues for novel aspects of virulence that will be explored in further studies.

INTRODUCTION

Bacterial regulatory RNAs represent a diverse class of regulators that operate at all layers of gene regulation, ranging from transcriptional initiation to protein translation (Papenfort and Vogel 2010; Harris et al. 2013). An emerging class of such regulators are from 40-500 nucleotides in length and are thus called small RNAs – sRNAs (Li et al. 2012).

Most sRNAs can be divided in the following four broad categories: (1) *cis*-acting RNAs; *trans*-acting RNAs that may either (2) modulate protein activity or (3) bind to mRNAs; and (4) Clustered Regularly Interspaced Short Palindromic Repeats – CRISPRs (Michaux et al. 2014b). The most studied bacterial sRNAs are the ones coded in *trans* which exert their cellular roles by base pairing with mRNA targets to attenuate, stop or activate their translation (Man et al. 2011; Papenfort and Vanderpool 2015). These sRNAs normally have more than one target since they only have limited complementarity with their cognate mRNAs (Han et al. 2013). Because of this partial complementarity, some of them may rely on the molecular chaperone Hfq to mediate their proper interaction with the cognate mRNAs by remodeling and stabilizing their structure, in addition to stimulating annealing (Vogel and Luisi 2011).

sRNAs that interact with proteins include the 6S RNA, which binds to the primary holoenzyme form of RNA polymerase and affects the expression of housekeeping genes under low nutrient conditions (Cavanagh and Wassarman 2014), and the sRNA CsrB, which is the major regulator of the protein CsrA - the effector of the complex network of the carbon storage regulatory (Csr) system controlling various virulence-related and metabolic phenotypes in several bacteria (Vakulskas et al. 2015).

Among the regulatory RNAs that act in *cis*, antisense sRNAs and riboswitches are the most important. Antisense sRNAs are transcribed from the DNA strand opposite their target gene on the bacterial chromosome, with which they have perfect complementarity (Thomason and Storz 2010). These RNAs can also be regulators of virulence and stress response in important pathogens (Gomez-Lozano et al. 2014b; Cho and Kim 2015). Riboswitches consist of mRNAs' regulatory segments, which alter their conformation in

response to the presence of a particular metabolite, usually causing the ribosome binding site in the cognate mRNA to be blocked or exposed (Mandal and Breaker 2004; Narberhaus et al. 2006).

CRISPR elements and their associated Cas proteins are considered the adaptive immunity system in prokaryotes that function via a mechanism of foreign DNA fragment (mainly bacteriophages and plasmids) incorporation into repeated arrays and subsequent utilization of transcripts of these inserts (known as spacers) as guide RNAs to cleave the cognate selfish element genome (Koonin and Wolf 2015).

Since sRNAs play versatile roles in the bacterial cell, a determined sRNA profile guarantees a quick and precise process of gene regulation and physiological adaptation to an ever-changing environment, which may be necessary for the establishment of a bacterial pathogenic lifestyle (Michaux et al. 2014b).

Many studies of novel sRNA identification rely on RNA sequencing (Li et al. 2013; Bilusic et al. 2014; Gomez-Lozano et al. 2014a), but because most sRNAs in bacterial transcriptomes correspond to a few overexpressed structural RNAs or products of mRNA degradation, only a portion of these RNAs is identified in laboratory approaches, requiring more expensive and time-consuming protocol adaptations (Gomez-Lozano et al. 2014a). For that reason, computational tools have become relevant, with ever-growing approaches for the discovery and characterization of regulatory RNAs (Cros et al. 2011; Livny 2012; Tesorero et al. 2013;Corredor and Murillo 2014).

In this work, we focused on the pathogenic bacterium *Actinobacillus pleuropneumoniae*, the causative agent of swine pleuropneumonia, a severe necrotic, fribrinous and hemorrhagic disease (Bossé et al. 2002; Krejci and Newberry 2011). *A. pleuropneumoniae* is a Gram-negative microaerophilic cocobacillus of the family *Pasteurellaceae*. The pathogenesis of pleuropneumonia is complex and involves many virulence factors, of which the Apx toxins, of the RTX family, are believed to play a central role (Frey 2011). Recently, *A. pleuropneumoniae hfq* mutants were shown to be defective in biofilm formation, displayed enhanced sensitivity to oxidative stress (Subashchandrabose et

al. 2013), and were attenuated in an alternative infection model (Pereira et al. 2015b). Hfq is an RNA-binding protein that facilitates the pairing of small RNAs with their target mRNAs and affects gene expression (Vogel and Luisi 2011). Although these findings indicate that *A. pleuropneumoniae* may rely on small RNAs to control aspects of its virulence, no studies on regulatory RNAs have been reported for this bacterium so far. In addition, little is known about the role of sRNAs in other *Pasteurellaceae*. The first study of these regulators in this family was performed on the human oral pathogen *Aggregatibacter actinomycetemcomitans*, in which three novel iron-regulated sRNA were identified (Amarasinghe et al. 2012), followed by the discovery of the iron-regulated HffR sRNA in *Haemophilus influenzae* (Santana et al. 2014). Thus, the main goal of this study was to establish a pipeline using free, web-accessible and user-friendly computational tools for the identification of regulatory RNAs in *Actinobacillus pleuropneumoniae* and other bacteria from the same family. Selected examples of predicted sRNAs were confirmed by Northern blotting and/or RT-PCR demonstrating the utility of the approach.

RESULTS

Computational prediction of regulatory RNAs in *Actinobacillus pleuropneumoniae*

For the discovery of novel regulatory RNAs in the *A. pleuropneumoniae* L20 genome, we organized a pipeline relying on the combination of four algorithms based on different approaches, such as the search for intergenic regions that are evolutionarily conserved (RNAz), that may contain transcriptional terminators (SIPHT), or may form stable secondary structures or characteristic motifs (INFERNAL), in addition to previously described sRNAs (BLASTn against Rfam), followed by experimental validation (Fig. 1). Because each program can generate a high number of candidates, the results obtained from each method were compared with one another to increase prediction accuracy. Sequences that were predicted by at least two different algorithms were considered to be sRNA candidates for further evaluation.

The algorithms used (RNAz, INFERNAL, SIPHT, and BLASTn against Rfam) predicted 215, 177, 44 and 108 genomic segments as putative regulatory RNAs, respectively. Many RNAs predicted by BLASTn/Rfam were tRNAs or rRNAs and were discarded from the analysis. The final number of RNA candidates considered after checking the intersection of the four results was 23, as shown by the Venn diagram in Figure 2A. They were named as Arrc01-23, from ***A****ctinobacillus pleuropneumoniae* **R**egulatory **R**NA **C**andidate.

Among the 23 candidates, eight (Arrc01, 03, 06, 10, 13, 15, 17 and 19) were predicted by at least three algorithms, with Arrc01and Arrc15 predicted by all four employed. All the Arrc *loci* are within intergenic regions and thus are not annotated in the *A. pleuropneumoniae* L20 genome, or any of the other complete genomes from this species available in the public databases. They are described in Table 1.

The RNA candidates identified can be classified in different categories, such as housekeeping regulatory RNAs, *trans*-acting sRNAs modulating protein activity, *trans*-acting sRNAs regulating mRNAs, and *cis*-acting RNAs, as will be presented in the next section. The present pipeline was not designed to detect CRISPRs, as these are not targets of the algorithms used. However, a separate search with the CRISPRFinder program (Grissa et al. 2007) detected a CRISPR element of 1503 bases, composed of 26 typical nearly-identical sequences of 28 bases each, all separated by spacers (25 in total) that are in average 32 bases long and surrounded by **C**RISPR-**as**sociated proteins, Cas (Fig. 2B).

Verification of the regulatory RNA candidates' expression

Prior to performing Northern blotting, the coding strand and expression of the sRNA candidates was verified by RT-PCR. The coding strand was determined by using only the forward or the reverse primer designed for each candidate, in the cDNA synthesis reaction. By doing so, only the reaction to which the primer capable of annealing to the sRNA was added would generate a cDNA product to be detected in the next PCR step of the protocol. In this analysis, we detected the expression of 17 of the 23 (74%) Arrcs (*data not shown*). It was not possible to detect the expression of Arrc03, 09, 12, 16, 18 and 22 in the conditions

evaluated. Then, to confirm the expression and relative abundance of the candidates, Northern blotting was employed. In every case, the Northern blot showed a discrete band with a size similar to the predicted sRNA, and in some situations, the blot also showed additional bands. The *A*. *pleuropneumoniae* ribosomal small RNA 5S was used as a positive control (Fig. 3, 4 and 5).

Housekeeping regulatory RNAs

Although more than half of all the RNAs predicted herein have not previously been described, the identity of some could be inferred by homology searches (BLASTn) against the main public databases. For example, Arrc06 is a widely distributed housekeeping RNA that is the functional RNA component of the Signal Recognition Particle (SRP) that delivers nascent peptides to their proper destination (Grotwinkel et al. 2014). Also very conserved and widespread, but with activities not related to the interaction with proteins, are the RNAs Arrc15 and Arrc23. Arrc15 is the ribozyme RNAseP, involved in processing tRNAs (Evans et al. 2006), and Arrc23 is a tmRNA, with dual tRNA-like and mRNA properties, which plays a central role in the process of recycling ribosomes stalled in aberrant mRNAs (Keiler and Ramadoss 2011). Both RNAseP and tmRNA use protein cofactors, which are also present in *A. pleuropneumoniae* genome. The genes *smp*B (small protein B, cofactor of tmRNA) and *rnp*A (protein C5, cofactor of RNAseP) are in the following position of *A. pleuropneumoniae* L20 genome sequence, respectively: 1006549-1007028 and 2172077-2172379. As would be expected, the expression of all of these RNAs was observed by Northern blotting (Fig. 3). No apparent differences in their level of expression were observed between the aerobic and anaerobic growth conditions.

Cis-acting RNAs

Our approach also predicted four *cis*-acting regulatory RNAs. Arrc03, Arrc13, Arrc19, and Arrc22 are homologous to the lysine, flavine mononucleotide (FMN), histidine, and molybdenum riboswitches, respectively. These annotations are consistent with the genome

localization of the candidates Arrc13, Arrc19, and Arrc22, as the first one is upstream to the gene *rib*D (riboflavin biosynthesis protein), the second is upstream to the gene *his*G (an ATP phosphoribosyl transferase involved in histidine biosynthesis), and the third is upstream to the gene *moa*A (molybdenum cofactor biosynthesis protein A). The annotation of Arrc03, however, remains unclear, since in its vicinity are the gene *hns* and the gene for a hypothetical protein which by BLASTn belongs to a family of Na+/H+ anti-porters (Pfam ID: pfam03553), not directly related to lysine biosynthesis, as expected.

Only the expression of the FMN and *his* riboswitches were observed by both RT-PCR and Northern blotting (Fig. 4). As expected for producing a coenzyme of the electron respiratory chain, the expression of FMN was more prominent during aerobic growth, while no apparent difference could be observed in the expression of the histidine riboswitch - probably because no nutritional stress was implicated in the growth conditions tested. These sRNA structures are in agreement with the fact that *cis*-regulatory elements usually include intrinsic attenuators (secondary structures shown in Supplemental Figure S1), frequently formed upon binding of the target molecule, thereby prematurely terminating transcription. Because riboswitches can be regulators at the transcription level (Henkin 2008), and the RNA extraction protocol was specific for purifying small molecules, the Northern blots of Arrc13 and Arrc19 showed a specific small band for the *cis*-element alone, instead of the entire unit containing the riboswitch and the regulated mRNA.

GcvB,6S and other putative trans-acting sRNAs

Among the 12 remaining regulatory RNAs whose expression was detected by RT-PCR, two are well-studied sRNAs, and ten are likely novel *trans-acting* RNAs. Arrc01 is the vastly studied GcvB, a major regulator of amino acid metabolism (Stauffer and Stauffer 2013), and Arrc10 is the global transcription regulator 6S RNA. Additionally, both Arrc20 and Arrc21 are homologous to sRNAs with broader targets. In the Rfam database, they belong to the RtT family, which in *Escherichia coli* was discovered as a RNA molecule liberated from the transcript of a tRNA operon and was implicated in cellular responses to face amino acid

limitations in the cell (Bosl and Kersten 1991). Their annotation is consistent with their location upstream of tRNA genes. Arrc20 is upstream of an Asn-tRNA gene and Arrc21 begins upstream of a Lys-tRNA gene.

With the exception of the Arrcs 02 and 07, all the other putative sRNAs showed consistent signals on Northern blots, including GcvB, the 6S and the RtTs (Fig. 5). While most sRNAs displayed single and specific bands, Arrc10 and Arrc17 exhibited additional shorter bands. With the exception of this additional shorter band observed for Arrc10 when the bacterium was grown in anaerobiosis, no other evident differences in any of the *trans* sRNAs could be observed when the bacterium was grown aerobically or anaerobically. Because of the abundance in palindromic sequences – which aided in their discovery –all the sRNAs are possibly able to form complex secondary structures composed of several hairpins, as predicted by RNAfold (Gruber et al. 2008).

RNA sequencing

The expression of the regulatory RNA candidates was also investigated by RNA sequencing (RNAseq) after bacterial growth under aerobic and anaerobic conditions. Because the RNAseq experiments resulted in a low number of reads (1106169 for aerobic and 1333114 for anaerobic growth) and no differential expression ($p<0.05$) between the two conditions was observed by Cuffdiff (Trapnell et al. 2012) - data not shown - the search for our RNA candidates was performed in an assembled transcriptome built after merging the aerobic and anaerobic reads. The resulting mapping files generated were uploaded to NCBI's SRA (*Sequence Read Archive*) under the experiment SRX810211. From that transcriptome we could confirm the expression of the RNAs Arrc01, 04, 06, 08, 10, 15, 20, 21 and 23.

Overall, from 23 predicted RNAs in *A. pleuropneumoniae*, the expression of 17 was confirmed under the conditions tested. From these, three are housekeeping regulatory RNAs (SRP – Arrc06, RNAseP – Arrc15, and tmRNA – Arrc23), two are *cis*-acting RNAs (FMN – Arrc13 and *his* – Arrc19 riboswitches), two are well-known *trans*-acting sRNAs (GcvB –

Arrc01 and 6S – Arrc10), ten are putative *trans*-acting sRNAs, of which two have Rfam homologues (RtTs – Arrc20 and Arrc21) and eight are novel sRNAs (Arrc02, 04, 05, 07, 08, 11, 14 and 17). All the candidates had putative Rho-independent terminator regions and promoter elements in the close upstream region of each designated gene, as predicted by BPROM (software Softberry, available at www.softberry.com, Supplemental Fig. S2). None of the sRNA genes described in this work had been previously annotated for the *A. pleuropneumoniae* genomes publicly available and represent, thus, an expansion in the understanding of the genome content of this microorganism.

The *trans*-acting sRNAs are potentially involved in an intricate network of gene regulation

To investigate the possible targets and roles of the aforementioned validated *trans*-acting sRNAs (Arrc01, 02, 04, 05, 07, 08, 11, 14, 17, 20 and 21), we performed a computational target prediction with TargetRNA2 (Kery et al. 2014), considering all the annotated ORFs in the *A. pleuropneumoniae* L20 genome. The interactions within the vicinity of the mRNA translational start site with the lowest energies and *p*-value below 0.05 were considered to indicate the best mRNA candidates. These targets are depicted in Supplemental Table S2. Following the search criteria established, TargetRNA2 predicted from 7 (for Arrc20) to 36 (for Arrc02) - with an average of 19.09 ± 8.53 - known genes whose mRNAs present great probability of binding the respective sRNA. Most of these targets are predicted to preferentially bind to conserved specific regions of the cognate sRNA (Fig. 5).

Because some of the target genes are common to more than one sRNA, many sRNAs are likely linked in an entangled and complex gene regulatory network (Fig. 6). Many candidates have the potential to control the translation of mRNAs directly involved in virulence. For example, the mRNA for the gene *apxII*A, encoding one of the Apx exotoxins, is predicted to be one of the targets of Arrc21 (RtT). Several sRNAs potentially bind mRNAs from genes whose products are involved in the intake and transport of iron from the host. Arrc05 may bind the RNA from the locus APL_0271, whose product is part of an iron

transport system. Likewise, Arrc14 may control the expression of a ferric permease (gene *afu*B2). The expression of different ferredoxins, encoded by the genes *napF*, *fdx* and locus APL_1678 are also potentially controlled by the sRNAs Arrc07, Arrc14 and Arrc20, respectively. The possible target of Arrc17, the mRNA for the gene *pbpB*, is involved in the resistance to penicillin. Two Arrc14 and one Arrc07 targets are related to cell adherence and biofilm formation; the first one may control the expression of tight adherence proteins coded by the genes *tad*E and *tad*D, and the latter may regulate the expression of the biofilm synthesis protein PgaA. Also, cell surface components involved with either adherence or protection are the possible targets of Arrc02. It may control the product of the locus APL_1273, which is a fimbrial biogenesis protein. Other sRNA candidates may control the expression of mRNA targets involved with the bacterial resistance to stressful conditions. The sRNA Arrc05 likely binds three of these mRNAs, transcribed from the genes *usp*A, *ost*A and *rec*J. These genes encode a universal stress protein, an organic solvent tolerance protein, and an exonuclease involved with DNA damage repair, respectively. The genes *rdg*C and *rad*A (likely targets of Arrc02 and Arrc11, respectively) also encode proteins that participate in processes of repair and recombination. The expression of four genes coding for heat shock proteins and/or chaperones may be controlled by Arrc07 (gene *grp*E), Arrc11 (gene *djl*A), and Arrc14 (genes *htp*G and *tor*D). Finally, some targets predicted for GcvB (Arrc01), highly involved in amino acid transport and metabolism, have already been described for this sRNA (Sharma et al. 2011), such as the genes *lrp*, *ilv*C, *ilv*E and *ser*A, all linked to amino acid biosynthesis.

Distribution of the sRNAs among *Pasteurellaceae*

Overall, 51 complete genomes available in Genbank from 15 different species of the *Pasteurellaceae* family were searched for similar RNAs sequences found in *A. pleuropneumoniae*, adopting a cutoff of 60% of identity and coverage. Among them, only *Mannheimia succinoproducens* and *Actinobacillus succinogenes* are not usually pathogenic. The distribution of the sequences of the regulatory RNA candidates described in this work

ranges from 100% (23 out of 23, for all *A. pleuropneumoniae* serotype references and all the Brazilian clinical isolates) to 13% (3 out of 23, for *Aggregatibacter actinomycetemcomitans*, *Aggregatibacter aphrophilus*, *Haemophilus parainfluenzae* and *Pasteurella multocida*). The results are shown in Supplemental Table S3. The sequence conservation and taxonomical dispersion among the 17 additional - apart from L20 - *A. pleuropneumoniae* strains (including 11 serotype references and six Brazilian clinical isolates) and another 14 *Pasteurellaceae* species (comprising 33 complete genomes) is also depicted in Table 1.

None of the *A. pleuropneumoniae* putative sRNAs are exclusive to the species, as all the sequences were found (100% of distribution) in the genome of *A. suis* 130Z, which is also a pig pathogen and closest species to *A. pleuropneumoniae*, as recently shown by phylogenomics (Naushad et al. 2015). The species of the genus *Mannheimia* are the next ones sharing the highest number of sRNA sequences with *A. pleuropneumoniae*. The strains of *Mannheimia varigena*, a bovine respiratory pathogen, present an average of 52% (12/23) of the sRNAs in common with *A. pleuropneumoniae*, while *Mannheimia haemolytica*, also a bovine respiratory pathogen, presents an average of 48% (11/23). Supplemental Figure S2 shows the sequence alignment for the novel *trans*-acting sRNAs described in this work.

Only three regulatory RNA candidates are well conserved among all the genomes analyzed: The RtT (Arrc21), the RNAseP (Arrc15) and the tmRNA (Arrc23) are present in all the *Pasteurellaceae* genomes available in at least 80% of the species. Additionally, sequences similar to GcvB are present next to the *gcv*A gene (or its homolog) identified in most of the *Pasteurellaceae* genomes analyzed (Fig. S2).

DISCUSSION

RNA molecules play a great variety of regulatory roles in all life domains, and increasing evidence shows they are implicated in virtually every aspect of cell metabolism (Waters and Storz 2009). This is especially important for bacterial pathogens, whose lifestyles require tight control of virulence gene expression and general stress responses (Papenfort and Vogel 2010; Bilusic et al. 2014). Several studies have revealed the existence

of sRNAs particularly involved in pathogenicity. As a result, an ever-growing library of virulence-related regulatory RNAs is being established in this relatively young field of life science. As recent specific examples, there are: the sRNA RhyB, which participates in the regulation of the production of siderophores in *Escherichia coli* (Porcheron and Dozois 2015); the *Staphylococcus aureus* sRNA teg49, whose lack hampers biofilm formation (Kim et al. 2014), the sRNA NrsZ, which modulates *Pseudomonas aeruginosa* motility (Wenner et al. 2014); and four novel sRNAs in *Enterococcus faecalis*, whose deletions affected bacterial virulence and stress tolerance when compared to the wild type strains (Michaux et al. 2014a). CRISPR elements are usually considered as the bacterial innate immune system to face mainly bacteriophage and conjugative plasmid invasions (Koonin and Wolf 2015), and increasing evidence show their differential expression during responses to stressful changes in the environment and during infection (Louwen et al. 2014).

For the *Pasteurellaceae* family, there have been few studies regarding sRNAs so far, to our knowledge. Although the relevance of sRNAs in *A. pleuropneumoniae* pathogenesis is suggested by previous work with mutants for the sRNA molecular chaperone Hfq (Subashchandrabose et al. 2013; Pereira et al. 2015b), their existence, abundance and possible roles were yet undescribed.

The discovery of novel regulatory RNAs has largely relied on RNAseq experiments, which may or may not be preceded by the co-immunoprecipitation of RNAs with the Hfq chaperone (Li et al. 2013; Bilusic et al. 2014; Gomez-Lozano et al. 2014a). Bioinformatics is also an important and less expensive tool to complement or replace the searches for regulatory RNAs in microorganisms and has been used with success (Tesorero et al. 2013), including the work with the *Pasteurellaceae* human pathogen *A. actinomycetemcomitans* (Amarasinghe et al. 2012). Although the computational prediction of sRNA candidates greatly diminishes the costs and time of initial experiments, the difficulties in handling algorithms generally limits their use by most biologists. To circumvent this predicament, in this work, we propose a pipeline using up-to-date programs, all hosted in free online and easy-to-use platforms, for the discovery of regulatory RNAs, using *A. pleuropneumoniae* as a model. We

chose the serovar 5 strain L20 as the basis for our study because it is considered to be highly virulent, and it was the first strain for which a complete closed genome was available (Foote et al. 2008).

As reviewed previously (Sridhar and Gunasekaran 2013), many algorithms following distinct parameters for the discovery of small regulatory RNAs have been created. Since single algorithm runs can result in a very high and unrealistic number of RNA candidates to logistically screen for, as was found in this study (see below), we used four of the available algorithms, each one with a different approach, and then compared the results in order to select candidates indicated by at least two of the methods employed. A similar strategy was successfully adopted and led to the discovery of seven novel small RNAs in *Streptococcus pyogenes* (Tesorero et al. 2013).

The algorithms RNAz, INFERNAL and BLASTn were all used through the platform RNAspace (Cros et al. 2011), and SIPHT was used through its own platform (Livny 2012). RNAz is a method of comparative genomics that searches for conserved genome fragments which present small RNA motifs, while evaluating the thermodynamic stability of their secondary structure (Gruber et al. 2010). Also a comparative model, INFERNAL takes the genome sequence used as the input to build consensus RNA secondary structure profiles, called covariance models, and uses them to search nucleic acid sequence databases for homologous RNAs (Nawrocki and Eddy 2013). The simple BLASTn approach was used for searching sRNAs that have already been described and deposited in Rfam, the greatest database for non-coding RNAs (Nawrocki et al. 2014). And finally, SIPHT basically identifies sRNA candidates by searching for conserved intergenic regions upstream of predicted intrinsic Rho-independent transcription terminators (Livny 2012). For our search with SIPHT, moderate stringent parameters were used because in a previous search, Livny et al.(2008) showed that those specific values were the ones that generated the lowest number of false sRNA candidates, taking into consideration the analyses of genomes of microorganisms whose sRNAs content had been previously studied. Because of that, SIPHT predicted about three times fewer candidates than the other three algorithms.

If not analyzed in combination, the four algorithms would have predicted 512 different RNA candidates (195 predicted exclusively by RNAz, 165 by INFERNAL, 32 by SIPHT and 97 by BLAST, plus 23 predicted by two or more programs), a number very far from the expected for a microorganism with a genome of approximately 2.2 Mb, such as *A. pleuropneumoniae*. As a comparison, in *E. coli*, whose genome is about twice as big as the *A. pleuropneumoniae*'s genome, around 80 sRNAs have already been described (Modi et al. 2011). We would therefore expect roughly half the regulatory RNAs in our model microorganism.

Here, we predicted 23 regulatory RNA candidates and observed the expression of 17 of them in aerobic and anaerobic conditions, mostly by RT-PCR and abundance by Northern blotting. We used aerobiosis as it is the typical atmosphere used in experiments involving *A. pleuropneumoniae* and anaerobiosis as, based on mutant (Baltes et al. 2005; Jacobsen et al. 2005; Buettner et al. 2009) or transcriptome (Deslandes et al. 2010; Klitgaard et al. 2012) studies, it is representative of the growth conditions found in necrotic lungs of pigs. As shown in Fig. 4 and Table 1, some trans-acting sRNA candidates, like ARRC05 and ARRC08, are located between genes within operons, which could allow them to be mistaken for sub-products of polycistronic mRNA maturation. However, features of their sequences and vicinity, like promoters and transcription terminators, indicate they are indeed sRNAs. Their functional characterization, which is underway, is a tougher task to perform, since producing mutants for these genes could interfere with the expression of the operon where they are located.

Only 9/23 RNAs(39%) were also identified by RNAseq, which could be explained by the low number of reads obtained and by sequencing errors inherent of the IonTorrent platform (Mardis 2013). It is possible that the use of a different new generation sequencing platform allowing a greater number of reads and/or accuracy would have identified the remaining 14 sRNAs. However, our results suggest that prediction of sRNA by software can be informative. For *E. coli*, the minimum of 2 million reads from an IlluminaHiSeq sequencer was necessary to cover (at least one time) 96% of each expectedly expressed ORF (Haas et

al. 2012). Then, it is still possible that some of the six undetected RNAs exist but are not expressed, or are expressed in low levels, in the conditions tested. Nevertheless, the pipeline proposed herein had an accuracy of almost 74%, slightly raising the 72% accuracy of the suggested SIPHT parameters alone (Livny et al. 2008), and a more than two-fold higher than the 31% accuracy of a similar approach used for *S. pyogenes*, in which 14 out of 45 RNA candidates predicted by the combination of three different algorithms had their expression confirmed (Tesorero et al. 2013). Since this is, to our knowledge, the first effort to discover regulatory RNAs in *A. pleuropneumoniae*, the transcripts observed in this work greatly contribute to the understanding of novel genes in this microorganism's genome and in the family's genomes as well. Even more remarkable is the fact that many of these RNAs have no homologues in the Rfam database, therefore raising the global repertoire of regulatory non-coding RNA. This is also the first time that a CRISPR element is described for *A. pleuropneumoniae*, although its features will be better studied in a future work. It remains to be elucidated if the system is still active in this bacterium, because despite having several genes for the typical Cas proteins, some essential expected proteins, such as the spacer-acquisition protein Cas2 (van der Oost et al. 2014), are absent.

As expected for a bacterial pathogen, the overall 23 sRNAs include regulatory molecules with important implications for the bacterial cell maintenance and pathogenesis. Among them, one of the most widespread and abundant (approx. 10,000 copies per cells in stationary phase) is the 6S RNA, a global regulator sRNA that reduces the expression of several sigma-70 dependent promoters, favoring the interaction of RNA polymerase with alternate sigma factors, such as RpoS in *Bacillus subtilis* (Cavanagh and Wassarman 2014), and has been implicated in the down-regulation of the expression of key pathways in response to changing stressful conditions and growth adaptation (Cavanagh et al. 2010; Cavanagh and Wassarman 2013). GcvB is also one of the most highly conserved Hfq-associated small RNAs in Gram-negative bacteria and was previously reported to regulate many genes involved in the transport an biosynthesis of oligopeptides and amino acids, such as the branched-chain amino acid (BCAA) transport system (Sharma et al. 2011; Stauffer

and Stauffer 2013). The BCAA biosynthesis and transport system is well studied in *A. pleuropneumoniae*. The presence of these amino acids is required for the survival of the bacterium and their lack is responsible for the expression of both genes for their own biosynthesis, and virulence-related genes, as demonstrated in pigs (Wagner and Mulks 2006; Subashchandrabose et al. 2009). GcvB is also known to regulate the PhoQ-PhoP two-component system, which is involved in magnesium homeostasis, pathogenicity, cell envelope composition and acid resistance in several bacterial species (Coornaert et al. 2013). Also involved in stress response, the sRNAs of the RtT family were first discovered as oligonucleotides released from the primary transcript of the tyrT and many other tRNA operons in *E. coli* during the tRNA processing. These sRNAs present modulatory effects on stringent response and are overproduced during cell contact with antibiotics (Bosl and Kersten 1991; Kohanski et al. 2007). The *cis*-acting RNA FMN has also been implicated in oxidative stress resistance in *Deinococcus radiodurans* (Yang et al. 2014).

A total of 12 *trans*-acting sRNAs were described, and eight of them are novel regulators whose influences in the bacterial cell are yet unknown. As it has been proposed before (Gruber and Sperandio 2015; Kim et al. 2015; Peng et al. 2015), computational analysis of the sRNAs' targets is a fine starting point towards understanding their physiological roles in cell. Here, we used TargetRNA2 as a target predictor for being the algorithm with the best correlation of targets predicted and actually confirmed, among the programs widely used for this purpose (Kery et al. 2014). Because many of the mRNAs predicted are potential targets of more than one sRNA, these regulators may share some of their targets, placing them in a characteristic entangled network of gene regulation (Modi et al. 2011). Our predictions are strongly corroborated by the fact that several of the targets predicted for the GcvB (Arrc01) sRNA had been shown for other microorganisms and are consistent with its role (Sharma et al. 2011).

Most sRNAs have as targets mRNAs directly implicated in either virulence or stress resistance. Although this was partially expected given the aforementioned phenotypes of reduced biofilm-forming capacity, sensitivity to oxidative stress (Subashchandrabose et al.

2013), and reduced virulence in an alternative infection model (Pereira et al. 2015b) displayed by *A. pleuropneumoniae* mutants for the *hfq* gene, new candidates for virulence determinants were defined by the present regulatory network. Two categories of potential targets that are part of extremely important systems for the establishment of *A. pleuropneumoniae* as a pig pathogen and thus, strong mutant candidates for live attenuated vaccines, are the iron-acquisition systems and Apx toxins. While the ability of *A. pleuropneumoniae* to overcome iron-restriction is essential to the pathogen permanency in the host, the pore-forming and cytolytic Apx toxins are directly related to the pathology of porcine pleuropneumonia (Jacques 2004; Frey 2011).

sRNAs are enriched in intergenic regions, which are longer and more conserved than the average intergenic regions in bacterial genomes (Tsai et al., 2015). For that reason, we performed a homology search of the sRNAs sequences identified for *A. pleuropneumoniae*, in the available complete genomes of other bacteria of the *Pasteurellaceae* family. Contrary to expectations, the novel sRNAs found in *A. pleuropneumoniae* are not widespread in *Pasteurellaceae*. The most widely distributed sequences in the family are the ones of housekeeping regulatory RNAs, whose functions are essential to the bacterium, such as the RNAseP (Arrc15) and tmRNA (Arrc23). *Actinobacillus suis* is the only pathogen included in this analysis that shares all the candidates' sequences with *A. pleuropneumoniae*, which can be explained by the evolutionary closeness of these species as recently reinforced by phylogenimocs (Naushad et al. 2015). However, unlike *A. pleuropneumoniae*, *A. suis* is not a primary pathogen but an opportunistic one that can also cause disease in pigs of all ages (Christensen and Bisgaard 2004; MacInnes et al. 2012). Besides their putative sRNAs, both pathogens have many virulence factors in common, which makes *A. suis* also capable of causing a hemorrhagic pleuropneumonia, though it most often causes septicemia and diseases such as arthritis and meningitis that are sequelae to septicemia (Ojha et al. 2010).

The other *Pasteurellaceae* with substantial numbers of RNA candidates' sequences in their genome in common with *A. pleuropneumoniae* are *M. haemolytica*, *M. varigena* and *H. parasuis*. The species of the genus *Mannheimia* are relatively close to *A. pleuropneumoniae*

as seen by the family's phylogenomics, though the same is not true for *H. parasuis*. Both *M. haemolytica* and *M. varigena* are occasional respiratory pathogens of cattle and pigs and a transcriptome study with *M. haemolytica* had already shown one putative sRNA in common with *A. pleuropneumoniae* (Reddy et al. 2012; Harhay et al. 2014). The similarities between *A. pleuropneumoniae* and *H. parasuis* may be explained by the fact that they share the same host. *H. parasuis* is one of the earliest and most prevalent colonizers of piglets in the farrowing house, a commensal of the respiratory tract, a common isolate from nasal secretions in pigs, and the cause of Glässers disease (Xu et al. 2011). It is expected that some of the sRNAs identified in this study might be important for the establishment of these bacteria within the porcine host and even for causing disease.

In conclusion, this work is one of few to base the initial search of bacterial regulatory RNAs exclusively on bioinformatics, an inexpensive and faster alternative to the most commonly used methods for sRNAs discovery. The success of the pipeline proposed herein is justified mainly by the use of up-to-date algorithms with different approaches easy to manipulate, showing that our strategy will be useful for the elucidation of novel regulatory RNAs in microbial genomes. Our findings are also a great step forward in the understanding of the coding potential of *A. pleuropneumoniae* and of the *Pasteurellaceae* family in general. Because many of the RNAs discovered are potentially involved in virulence, it is possible that mutants for those sRNAs can be effectively used as attenuated vaccines, providing new horizon for further studies.

MATERIALS AND METHODS

An overview of the computational and experimental strategies used to identify and characterize novel sRNAs in *A. pleuropneumoniae* are summarized in Figure 1 and explained with details below.

Bacterial strains and RNA extraction

The experiments were conducted with the *A. pleuropneumoniae* serotype 5 reference strain

L20. The strain was grown in Brain Heart Infusion (BHI) supplemented with Nicotinamide Adenine Dinucleotide (NAD - 10 µg ml^{-1}) at 37 °C until early stationary phase (8h) under aerobic (5% CO_2) and anaerobic conditions (anaerobic jar with Oxoid Anaerogen Sachet, Thermo Scientific). Total RNA extraction was performed by cell disrupting using the Lysing Matrix B tubes (MP Biomedicals, USA), followed by the procedures of the miRNeasy Mini Kit (Qiagen, Netherlands), according to the manufacturer's instructions. After extraction, the concentration and purity of the RNA was determined by Nanodrop and by 2100 Bioanalyzer (Agilent Technologies). The resulting total RNA (purity 1.8-1.9, A260/A280 ratio) was treated with one unit of RQ1 DNAse (Promega, USA) per µg of nucleic acid, and incubated for 60 min at 37 ºC.

In silico identification of sRNAs

The computational searches for discovery of sRNAs were performed with the genome of *A. pleuropneumoniae* strain L20 (Genbank access NC_009053). Four algorithms with distinct approaches were employed to increase prediction accuracy: RNAz (Gruber et al. 2010), INFERNAL (Nawrocki and Eddy 2013), SIPHT (Livny 2012) and BLASTn on Rfam (Nawrocki et al. 2014). The predictions with BLASTn, RNAz and INFERNAL were performed through the RNAspace platform (Cros et al. 2011), available at http://www.rnaspace.org/. The BLASTn approach consisted in performing a homology search against all the regulatory RNAs sequences available from the Rfam 10.0 database. The RNAz run was made with the default values, i.e., probability cutoff: 0.7, slice alignments longer than: 300, window size: 200, step size: 50. Genome-wide sequence alignment of *A. pleuropneumonie* L20 was made with the annotated genomes of *A. pleuropneumoniae* JL03 and AP76, *Haemophilus ducreyi* 35000HP and *H. influenza* PittEE (Genbank accesses NC_010278, NC_010939, NC_002940 and CP000671, respectively) through BLAST, filtering low complexity regions in both strands of query sequence and adopting an E-value threshold of 0.001, which are the default values in RNAspace. CG-seq was used for sequence aggregation, also using the RNAspace default parameters, score lambda parameter: 1, minimal and maximal length of a

conserved region: 30 and 500, minimum and maximum identity threshold: 60 and 100. For the prediction with INFERNAL, the query genome sequence was used to build a covariance model (CM), then used to search homologous RNAs sequences on the Rfam 10.0 database. For the SIPHT run, whose platform is available at http://newbio.cs.wisc.edu/sRNA/, the moderate stringency parameters suggested by the author were used, which are as follows; maximum E-value: 1E-15, minimum TransTerm confidence value: 87%, maximum RNAMotif score: -9, FindTerm scores: -10, and minimum and maximum lengths of predicted loci: 50 and 500 (Livny et al. 2008). Finally, the resulting RNA lists of each algorithm were compared against each other with BLASTn, and the sequences that were predicted by at least two of the four methods applied were considered as sRNA candidates. For the identification of CRISPRs in the genome of the *A. pleuropneumoniae* L20, the standard definitions of the software CRISPRFinder (Grissa et al. 2007) were used.

Taxonomical dispersion of the *A. pleuropneumoniae*'s sRNAs in the *Pasteurellaceae* family

The sequences of the final sRNA candidates were also searched by BLASTn in the other 11 serotype reference genomes available of *A. pleuropneumoniae* and in the genomes of six Brazilian clinical isolates (Pereira et al. 2015a) to investigate their distribution among the species. In addition, 33 genomes of 14 other *Pasteurellaceae* species, whose complete genomes are available from Genbank, were also searched (Supplemental Table S4).

Reverse Transcription Polymerase Chain Reaction (RT-PCR)

To identify the coding strand of each predicted regulatory RNA candidate, RT-PCR was performed. For cDNA synthesis, the ImProm-II Reverse Transcription System (Promega, Madison, USA) was used, according to the manufacturer's instructions, using either the forward or the reverse primer designed for each sRNA candidate separately (Supplemental Table S1). The cDNA reaction (20 µL) was further used in a PCR reaction using the primer pair for each sRNA. The PCR reaction was performed with 1 U of GoTaq DNA polymerase

(Promega, Madison, USA) in a final volume of 50 µL of enzyme buffer containing 1.5 mM $MgCl_2$, 0.2 mM of each dNTP and 0.2 µM of each primer in a thermal cycler Mastercycler pro (Eppendorf, Germany). The samples were initially denatured at 94 °C for 2 min, followed by 35 reaction cycles (94 °C for 1 min, 55 °C for 1 min and 72 °C for 30 s) and a final extension step at 72 °C for 5 min. The amplicons generated were analyzed after eletrophoresis in a 2.0% agarose gel. As a positive amplification control, 50 ng of *A. pleuropneumoniae* L20 genomic DNA was used as template and as negative control, a reaction with DNA-free total RNA that was not subjected to the reaction of cDNA synthesis was used.

Northern Blotting

Total RNA (10 µg) was run on a 10% TBE-urea gel and transferred to a Brightstar Plus nylon membrane (Applied Biosystems, USA). Hybridization was conducted with the DIG High Prime DNA Labeling and Detection Starter kit II (Roche, Switzerland), according to the manufacturer's instructions. The oligonucleotides pairs used for the RT-PCR reaction (SupplementalTable S1) were also used to construct 126 ± 27 bp digoxigenin-marked probes designed for the inner parts of the sRNA candidates' sequences with the PCR DIG probe synthesis kit (Roche, Switzerland). As hybridization controls we used dot blots with total DNA and all membranes used were hybridized with probes for the rRNA 5S, using the primer pair APP5SF/APP5SR (Supplemental Table S1).

RNA sequencing, reads mapping and transcriptome assembly

Total RNA was extracted from *A.pleuropneumoniae* MIDG2331, a clinical isolate from the UK, as described above and was treated with MICROBExpress Kit (Life Technologies, USA) for ribosomal RNA removal. The cDNA library construction for both aerobic and anaerobic conditions was carried out using the Ion Total RNA-Seq Kit v2 (Life Technologies, USA) according to manufacturer's protocols. Samples were loaded onto a 318 chip and sequenced on Ion torrent-PGM (Life technologies, USA) using default parameters (single-end, forward sequencing). The sequenced reads were mapped to the *A. pleuropneumoniae* L20 reference

strain genome using Burrows-Wheeler Aligner (BWA-MEM algorithm, default parameters) version 0.7.10 (Li and Durbin 2009). The resulting bam files were uploaded in NCBI-Short Read Archive (SRA) under the experiment access SRX810211. Transcriptome assembly was made with Cufflinks version 2.2.1 (Trapnell et al. 2012).

Investigation of putative mRNA targets

The potential mRNAs targets of the putative novel *trans*-acting RNA (sRNA) candidates identified were searched for every annotated gene on the *A. pleuropneumoniae* L20 genome. Searches were performed with the software TargetRNA2 (Kery et al. 2014), considering the conservation (compared to every sequenced replicon available in GenBank) and accessibility of each sRNA given as input, structural accessibility of the mRNA and potential interactions preceded by a hybridization seed around the translation start site, from 80 nucleotides upstream to 20 nucleotides downstream of it. Only target interactions with a *p*-value less or equal to 0.05 were reported. Potential mRNA targets shared by the sRNAs were identified and used as the basis for the design of a regulatory network comprising each sRNA and their putative targets.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGEMENTS

The authors thank CNPq (407849/2012-2), FAPEMIG (APQ-00232-13), CAPES/PROEX and BBSRC (BB/K021109/1, BB/G018553 and BB/M023052/1) for their financial support.

REFERENCES

Amarasinghe JJ, Connell TD, Scannapieco FA, Haase EM. 2012. Novel iron-regulated and Fur-regulated small regulatory RNAs in *Aggregatibacter actinomycetemcomitans*. *Mol Oral Microbiol* **27**: 327-349.

Baltes N, N'Diaye M, Jacobsen ID, Maas A, Buettner FF, Gerlach GF. 2005. Deletion of the anaerobic regulator HlyX causes reduced colonization and persistence of *Actinobacillus*

pleuropneumoniae in the porcine respiratory tract. *Infect Immun* **73**: 4614-4619.
Bilusic I, Popitsch N, Rescheneder P, Schroeder R, Lybecker M. 2014. Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. *RNA Biol* **11**: 641-654.
Bosl M, Kersten H. 1991. A novel RNA product of the *tyr*T operon of *Escherichia coli*. *Nucleic Acids Res* **19**: 5863-5870.
Bossé JT, Janson H, Sheehan BJ, Beddek AJ, Rycroft AN, Kroll JS, Langford PR. 2002. *Actinobacillus pleuropneumoniae*: pathobiology and pathogenesis of infection. *Microb Infect* **4**: 225-235.
Buettner FF, Bendalla IM, Bossé JT, Meens J, Nash JH, Hartig E, Langford PR, Gerlach GF. 2009. Analysis of the *Actinobacillus pleuropneumoniae* HlyX (FNR) regulon and identification of iron-regulated protein B as an essential virulence factor. *Proteomics* **9**: 2383-2398.
Cavanagh AT, Chandrangsu P, Wassarman KM. 2010. 6S RNA regulation of *rel*A alters ppGpp levels in early stationary phase. *Microbiology* **156**: 3791-3800.
Cavanagh AT, Wassarman KM. 2013. 6S-1 RNA function leads to a delay in sporulation in *Bacillus subtilis*. *J Bacteriol* **195**: 2079-2086.
Cavanagh AT, Wassarman KM. 2014. 6S RNA, a global regulator of transcription in *Escherichia coli*, *Bacillus subtilis*, and beyond. *Annu Rev Microbiol* **68**: 45-60.
Cho KH, Kim JH. 2015. Cis-encoded non-coding antisense RNAs in streptococci and other low GC Gram (+) bacterial pathogens. *Front Genet* **6**: 110.
Christensen H, Bisgaard M. 2004. Revised definition of *Actinobacillus sensu stricto* isolated from animals. A review with special emphasis on diagnosis. *Vet Microbiol* **99**: 13-30.
Coornaert A, Chiaruttini C, Springer M, Guillier M. 2013. Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB. *PLoS Genet* **9**: e1003156.
Corredor M, Murillo O. 2014. Identification of small non-coding RNAs in bacterial genome annotation using databases and computational approaches. *Adv Intel Syst Comput* **232**: 295-300.
Cros MJ, de Monte A, Mariette J, Bardou P, Grenier-Boley B, Gautheret D, Touzet H, Gaspin C. 2011. RNAspace.org: An integrated environment for the prediction, annotation, and analysis of ncRNA. *RNA* **17**: 1947-1956.
Deslandes V, Denicourt M, Girard C, Harel J, Nash JH, Jacques M. 2010. Transcriptional profiling of *Actinobacillus pleuropneumoniae* during the acute phase of a natural infection in pigs. *BMC Genomics* **11**: 98.
Evans D, Marquez SM, Pace NR. 2006. RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* **31**: 333-341.
Foote SJ, Bossé JT, Bouevitch AB, Langford PR, Young NM, Nash JH. 2008. The complete genome sequence of *Actinobacillus pleuropneumoniae* L20 (serotype 5b). *J Bacteriol* **190**: 1495-1496.
Frey J. 2011. The role of RTX toxins in host specificity of animal pathogenic *Pasteurellaceae*. *Vet Microbiol* **153**: 51-58.
Gomez-Lozano M, Marvig RL, Molin S, Long KS. 2014a. Identification of bacterial small RNAs by RNA sequencing. *Methods Mol Biol* **1149**: 433-456.
Gomez-Lozano M, Marvig RL, Tulstrup MV, Molin S. 2014b. Expression of antisense small RNAs in response to stress in *Pseudomonas aeruginosa*. *BMC Genomics* **15**: 783.
Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **35**: W52-57.
Grotwinkel JT, Wild K, Segnitz B, Sinning I. 2014. SRP RNA remodeling by SRP68 explains its role in protein translocation. *Science* **344**: 101-104.
Gruber AR, Findeiss S, Washietl S, Hofacker IL, Stadler PF. 2010. RNAz 2.0: improved noncoding RNA detection. *Pac Symp Biocomput* **15**: 69-79.
Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL. 2008. The Vienna RNA websuite. *Nucleic Acids Res* **36**: W70-74.
Gruber CC, Sperandio V. 2015. Global analysis of posttranscriptional regulation by GlmY and GlmZ in enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **83**: 1286-1295.
Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J. 2012. How deep is deep enough for RNA-Seq

profiling of bacterial transcriptomes? *BMC Genomics* **13**: 734.
Han Y, Liu L, Fang N, Yang R, Zhou D. 2013. Regulation of pathogenicity by noncoding RNAs in bacteria. *Future Microbiol* **8**: 579-591.
Harhay GP, Murray RW, Lubbers B, Griffin D, Koren S, Phillippy AM, Harhay DM, Bono J, Clawson ML, Heaton MP et al. 2014. Complete closed genome sequences of four *Mannheimia varigena* isolates from cattle with shipping fever. *Genome Announc* **2**: e00088-14.
Harris JF, Micheva-Viteva S, Li N, Hong-Geller E. 2013. Small RNA-mediated regulation of host-pathogen interactions. *Virulence* **4**:785-798
Henkin TM. 2008. Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev* **22**: 3383-3390.
Jacobsen I, Hennig-Pauka I, Baltes N, Trost M, Gerlach GF. 2005. Enzymes involved in anaerobic respiration appear to play a role in *Actinobacillus pleuropneumoniae* virulence. *Infect Immun* **73**: 226-234.
Jacques M. 2004. Surface polysaccharides and iron-uptake systems of *Actinobacillus pleuropneumoniae*. *Can J Vet Res* **68**: 81-85.
Keiler KC, Ramadoss NS. 2011. Bifunctional transfer-messenger RNA. *Biochimie* **93**: 1993-1997.
Kery MB, Feldman M, Livny J, Tjaden B. 2014. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res* **42**: W124-129.
Kim S, Reyes D, Beaume M, Francois P, Cheung A. 2014. Contribution of teg49 small RNA in the 5' upstream transcriptional region of *sar*A to virulence in *Staphylococcus aureus*. *Infect Immun* **82**: 4369-4379.
Kim T, Bak G, Lee J, Kim KS. 2015. Systematic analysis of the role of bacterial Hfq-interacting sRNAs in the response to antibiotics. *J Antimicrob Chemother* **70**:1659-1668
Klitgaard K, Friis C, Jensen TK, Angen O, Boye M. 2012. Transcriptional portrait of *Actinobacillus pleuropneumoniae* during acute disease-potential strategies for survival and persistence in the host. *PloS One* **7**: e35549.
Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, Collins JJ. 2007. A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* **130**: 797-810.
Koonin EV, Wolf YI. 2015. Evolution of the CRISPR-Cas adaptive immunity systems in prokaryotes: models and observations on virus-host coevolution. *Mol Biosyst* **11**: 20-27.
Krejci R, Newberry J. 2011. Pleuropneumonia in pigs - its importance and prevention. *Int Pigs Topics* **26**: 15-17.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754-1760.
Li SK, Ng PK, Qin H, Lau JK, Lau JP, Tsui SK, Chan TF, Lau TC. 2013. Identification of small RNAs in *Mycobacterium smegmatis* using heterologous Hfq. *RNA* **19**: 74-84.
Li W, Ying X, Lu Q, Chen L. 2012. Predicting sRNAs and their targets in bacteria. *Genomics Proteomics Bioinf* **10**: 276-284.
Livny J. 2012. Bioinformatic discovery of bacterial regulatory RNAs using SIPHT. *Methods Mol Biol* **905**: 3-14.
Livny J, Teonadi H, Livny M, Waldor MK. 2008. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS One* **3**: e3197.
Louwen R, Staals RH, Endtz HP, van Baarlen P, van der Oost J. 2014. The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol Mol Biol Rev* **78**: 74-88.
MacInnes JI, Mackinnon J, Bujold AR, Ziebell K, Kropinski AM, Nash JH. 2012. Complete genome sequence of *Actinobacillus suis* H91-0380, a virulent serotype O2 strain. *J Bacteriol* **194**: 6686-6687.
Man S, Cheng R, Miao C, Gong Q, Gu Y, Lu X, Han F, Yu W. 2011. Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res* **39**: e50.
Mandal M, Breaker RR. 2004. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* **5**: 451-463.
Mardis ER. 2013. Next-generation sequencing platforms. *Annu Rev Analytical Chem* **6**: 287-303.
Michaux C, Hartke A, Martini C, Reiss S, Albrecht D, Budin-Verneuil A, Sanguinetti M, Engelmann S, Hain T, Verneuil N et al. 2014a. Involvement of *Enterococcus faecalis* small

RNAs in stress response and virulence. *Infect Immun* **82**: 3599-3611.
Michaux C, Verneuil N, Hartke A, Giard JC. 2014b. Physiological roles of small RNA molecules. *Microbiology* **160**: 1007-1019.
Modi SR, Camacho DM, Kohanski MA, Walker GC, Collins JJ. 2011. Functional characterization of bacterial sRNAs using a network biology approach. *PNAS* **108**: 15522-15527.
Narberhaus F, Waldminghaus T, Chowdhury S. 2006. RNA thermometers. *FEMS Microbiol Rev* **30**: 3-16.
Naushad S, Adeolu M, Goel N, Khadka B, Al-Dahwi A, Gupta RS. 2015. Phylogenomic and molecular demarcation of the core members of the polyphyletic *Pasteurellaceae* genera *Actinobacillus*, *Haemophilus*, and *Pasteurella*. *Int J Genomics* **2015**: 198560.
Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J et al. 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res* **42**: D130-D137.
Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**: 2933-2935.
Ojha S, Lacouture S, Gottschalk M, MacInnes JI. 2010. Characterization of colonization-deficient mutants of *Actinobacillus suis*. *Vet Microbiol* **140**: 122-130.
Papenfort K, Vanderpool CK. 2015. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev* **39**: 362-378.
Papenfort K, Vogel J. 2010. Regulatory RNA in bacterial pathogens. *Cell Host Microbe* **8**: 116-127.
Peng X, Dong H, Wu Q. 2015. A new cis-encoded sRNA, BsrH, regulating the expression of hemH gene in *Brucella abortus* 2308. *FEMS Microbiol Lett* **362**: 1-7.
Pereira MF, Rossi CC, de Carvalho FM, de Almeida LG, Souza RC, de Vasconcelos AT, Bazzolli DM. 2015a. Draft Genome Sequences of Six *Actinobacillus pleuropneumoniae* Serotype 8 Brazilian Clinical Isolates: Insight into New Applications. *Genome Announc* **3**: e01585-14.
Pereira MF, Rossi CC, Queiroz MV, Martins GF, Isaac C, Bosse JT, Li Y, Wren BW, Terra VS, Cuccui J et al. 2015b. *Galleria mellonella* is an effective model to study *Actinobacillus pleuropneumoniae* infection. *Microbiology* **161**: 387-400.
Porcheron G, Dozois CM. 2015. Interplay between iron homeostasis and virulence: Fur and RyhB as major regulators of bacterial pathogenicity. *Vet Microbiol* **179**: 2-14.
Reddy JS, Kumar R, Watt JM, Lawrence ML, Burgess SC, Nanduri B. 2012. Transcriptome profile of a bovine respiratory disease pathogen: *Mannheimia haemolytica* PHL213. *BMC Bioinformatics* **13 Suppl 15**: S4.
Santana EA, Harrison A, Zhang X, Baker BD, Kelly BJ, White P, Liu Y, Munson RS, Jr. 2014. HrrF is the Fur-regulated small RNA in nontypeable *Haemophilus influenzae*. *PLoS One* **9**: e105644.
Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf HJ, Hinton JC, Vogel J. 2011. Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* **81**: 1144-1165.
Sridhar J, Gunasekaran P. 2013. Computational small RNA prediction in bacteria. *Bioinformatics Biol Insights* **7**: 83-95.
Stauffer LT, Stauffer GV. 2013. Multiple roles for the sRNA GcvB in the regulation of Slp levels in *Escherichia coli*. *ISRN Bacteriol* **2013**: 918106.
Subashchandrabose S, Leveque RM, Kirkwood RN, Kiupel M, Mulks MH. 2013. The RNA chaperone Hfq promotes fitness of *Actinobacillus pleuropneumoniae* during porcine pleuropneumonia. *Infection Immun* **81**: 2952-2961.
Subashchandrabose S, LeVeque RM, Wagner TK, Kirkwood RN, Kiupel M, Mulks MH. 2009. Branched-chain amino acids are required for the survival and virulence of *Actinobacillus pleuropneumoniae* in swine. *Infection Immun* **77**: 4925-4933.
Tesorero RA, Yu N, Wright JO, Svencionis JP, Cheng Q, Kim JH, Cho KH. 2013. Novel regulatory small RNAs in *Streptococcus pyogenes*. *PLoS One* **8**: e64021.
Thomason MK, Storz G. 2010. Bacterial antisense RNAs: how many are there, and what are they doing? *Annu Rev Genet* **44**: 167-188.
Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL,

Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**: 562-578.

Tsai CH, Liao R, Chou B, Palumbo M, Contreras LM. 2015. Genome-wide analyses in bacteria show small-RNA enrichment for long and conserved intergenic regions. *J Bacteriol* **197**: 40-50

Vakulskas CA, Potts AH, Babitzke P, Ahmer BM, Romeo T. 2015. Regulation of Bacterial Virulence by Csr (Rsm) Systems. *Microbiol Mol Biol Rev* **79**: 193-224.

van der Oost J, Westra ER, Jackson RN, Wiedenheft B. 2014. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* **12**: 479-492.

Vogel J, Luisi BF. 2011. Hfq and its constellation of RNA. *Nat Rev Microbiol* **9**: 578-589.

Wagner TK, Mulks MH. 2006. A subset of *Actinobacillus pleuropneumoniae* in vivo induced promoters respond to branched-chain amino acid limitation. *FEMS Immunol Med Mic* **48**: 192-204.

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

Wenner N, Maes A, Cotado-Sampayo M, Lapouge K. 2014. NrsZ: a novel, processed, nitrogen-dependent, small non-coding RNA that regulates *Pseudomonas aeruginosa* PAO1 virulence. *Environ Microbiol* **16**: 1053-1068.

Xu Z, Yue M, Zhou R, Jin Q, Fan Y, Bei W, Chen H. 2011. Genomic characterization of *Haemophilus parasuis* SH0165, a highly virulent strain of serovar 5 prevalent in China. *PLoS One***6**: e19631.

Yang P, Chen Z, Shan Z, Ding X, Liu L, Guo J. 2014. Effects of FMN riboswitch on antioxidant activity in *Deinococcus radiodurans* under H_2O_2 stress. *Microbiol Res* **169**: 411-416.

Table 1. Regulatory RNAs predicted *in silico* for *Actinobacillus pleuropneumoniae* in this work.

ID	Rfam (E-value)	Genome position[a]	Size	Left gene	Right gene	Strand	Method[b]	RT-PCR, Northern blot, RNAseq detection	Taxonomical dispersion[c]
Housekeeping Regulatory RNA									
Arrc06	SRP (4E-42)	925791-925977	187	*ispF*	APL_0804	+	B, Z, I	Yes, yes, yes	18/3
Arrc15	RNAseP (0.0)	599910-600092	182	APL_0525	APL_0526	+	B, S, Z, I	Yes, yes, yes	18/14
Arrc23	tmRNA (1E-179)	2059595-2059995	365	*glpT*	APL_1836	+	B, I	Yes, yes, yes	18/14
cis-acting RNAs									
Arrc03	Lys riboswitch (9E-82)	516179-516425	247	APL_0448	APL_0449	+	B, Z, I	No, no, no	18/4
Arrc13	FMN riboswitch (9E-57)	435489-435614	126	*glpC*	*ribD*	+	B, Z, I	Yes, yes, no	18/5
Arrc19	His riboswitch (3E-12)	2246012-2246156	145	*purC*	*hisG*	+	B, S, I	Yes, yes, no	18/1
Arrc22	Mo riboswitch (3E-06)	790908-791095	188	APL_0689	*moaC*	+	B, I	No, no, no	18/1
trans-acting sRNAs									
Arrc01	GcvB (1E-36)	146327-146527	201	*engC*	*gcvA*	-	B, S, Z, I	Yes, yes, yes	18/2
Arrc02	nd*	413161-413400	240	*gprE*	APL_0368	+	S, Z	Yes, no, no	18/2
Arrc04	nd	727911-728110	200	*hsdR3*	$tRNA^{Met}$	+	S, Z	Yes, yes, yes	18/1
Arrc05	nd	789422-789609	188	*torZ*	*torY*	-	Z, I	Yes, yes, no	18/1
Arrc07	nd	1304301-1304515	215	APL_1127	APL_1128	+	Z, I	Yes, no, no	18/5
Arrc08	nd	1950746-1951015	270	*rplA*	*rplJ*	+	S, Z	Yes, yes, yes	18/4
Arrc09	nd	2136337-2136497	161	*dnaJ*	*dnaK*	+	S, Z	No, no, no	18/2
Arrc10	6S (4E-87)	128740-128984	245	APL_0109	*zapA*	+	B, Z, I	Yes, yes, yes	18/3
Arrc11	nd	208880-209050	171	*sixA*	*dus*	+	S, Z	Yes, yes, no	18/1
Arrc12	nd	246118-246362	245	APL_0223	APL_0224	-	S, Z	No, no, no	18/4
Arrc14	nd	446936-447102	167	APL_0392	*leuA*	+	S, Z	Yes, yes, no	18/1
Arrc16	nd	1691444-1691630	187	APL_1476	APL_1477	+	S, Z	No, no, no	18/4
Arrc17	nd	2007501-2007634	134	*rpmJ*	*rpmM*	+	B, Z, I	Yes, yes, no	18/9
Arrc18	nd	2094311-2094468	158	APL_1863	APL_1864	+	S, Z	No, no, no	18/1
Arrc20	RtT (6E-08)	346142-346419	278	*moeB*	$tRNA^{Asn}$	+	B, Z	Yes, yes, yes	18/3
Arrc21	RtT (8E-07)	2041642-2041817	177	$tRNA^{Lys}$	*acpP*	-	B, Z	Yes, yes, no	18/12

[a]Coordinates and ORF names are given in relation to the genome of the *A. pleuropneumoniae* L20.
[b]Methods: "B" corresponds to the BLASTn algorithm over Rfam, "S" to SIPHT, "Z" to RNAz and "I" to Infernal.
[c]Sequence conservation and taxonomical dispersion were analyzed in 17 additional *A. pleuropneumoniae* strains (including 11 serotype references and six clinical isolates) and another 14 *Pasteurellaceae* species (comprising 33 complete genomes). Numbers denote the amount of bacterial strains in which Arrc homologs have been found in the respective order: *A. pleuropneumoniae* / Other *Pasteurellaceae*. Homologs were detected by BLAST using a 60% coverage and identity and an E-value of 0.001 as the thresholds.
*nd: sequences not detected in the Rfam database

FIGURE LEGENDS

FIGURE 1. A strategy for the search and characterization of regulatory RNAs in *A. pleuropneumoniae* (App). Black rectangles represent the tools, both computational and laboratorial, for the prediction, validation and characterization of non-coding regulatory RNA in *A. pleuropneumoniae*. White rectangles show inputs and grey rectangles show the outputs (and final objectives) in the workflow, represented by the arrows.

FIGURE 2. Computational prediction of small regulatory RNAs in *A.pleuropneumoniae*. (**A**) The sRNA candidates were defined as those predicted by at least two out of four (BLAST/Rfam, SIPHT, Infernal and RNAz) algorithms with different approaches. As shown by the Venn diagram's intersections, the number of final candidates selected adopting this criterion was 23. (**B**) A CRISPR locus was also searched by using the software

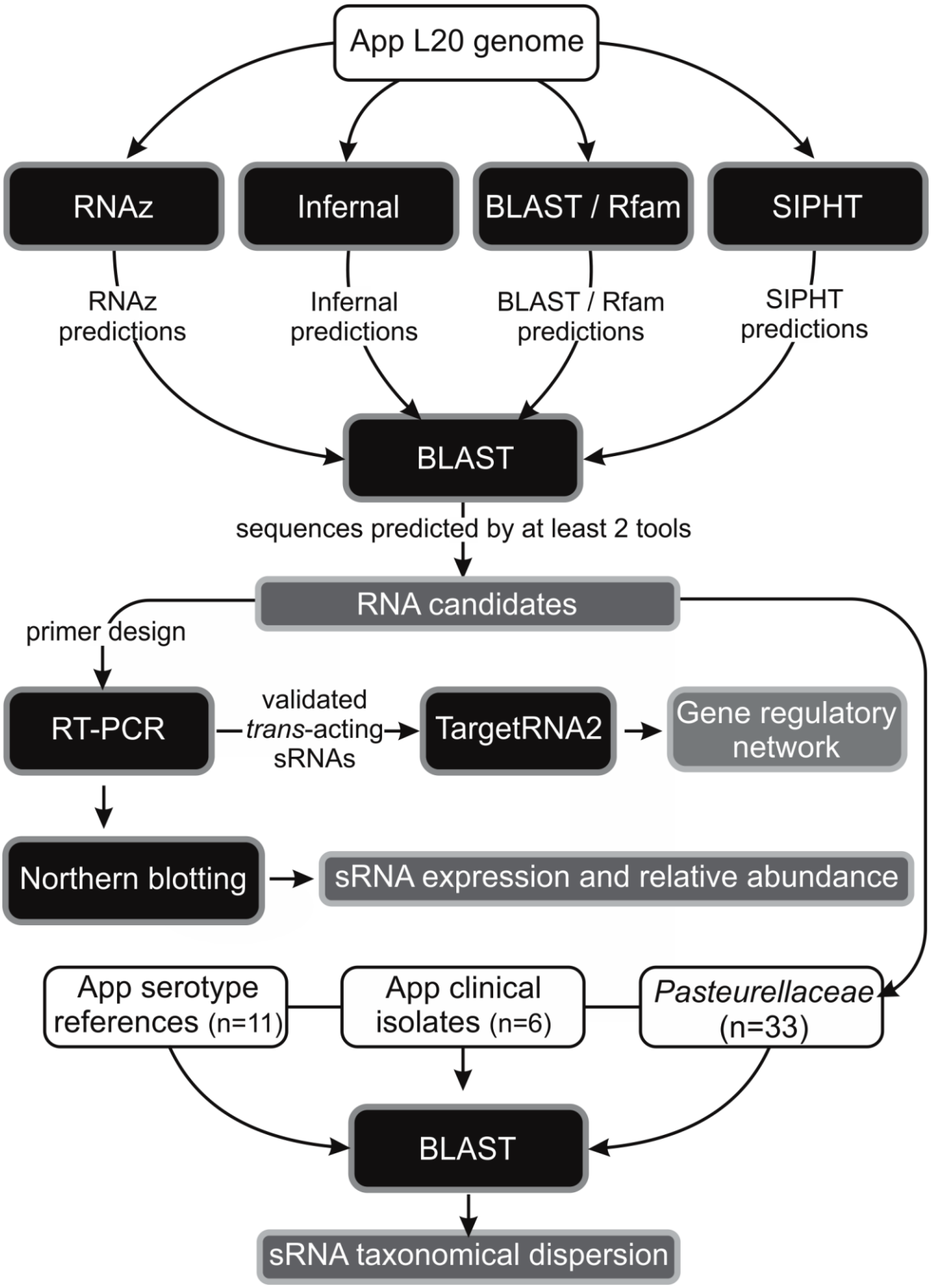
CRISPRFinder. For the *A. pleuropneumoniae* reference strain L20, a total of 26 direct repeats, separated by 25 spacers (grey) and surrounded by CRISPR-associated genes (arrows), were found.

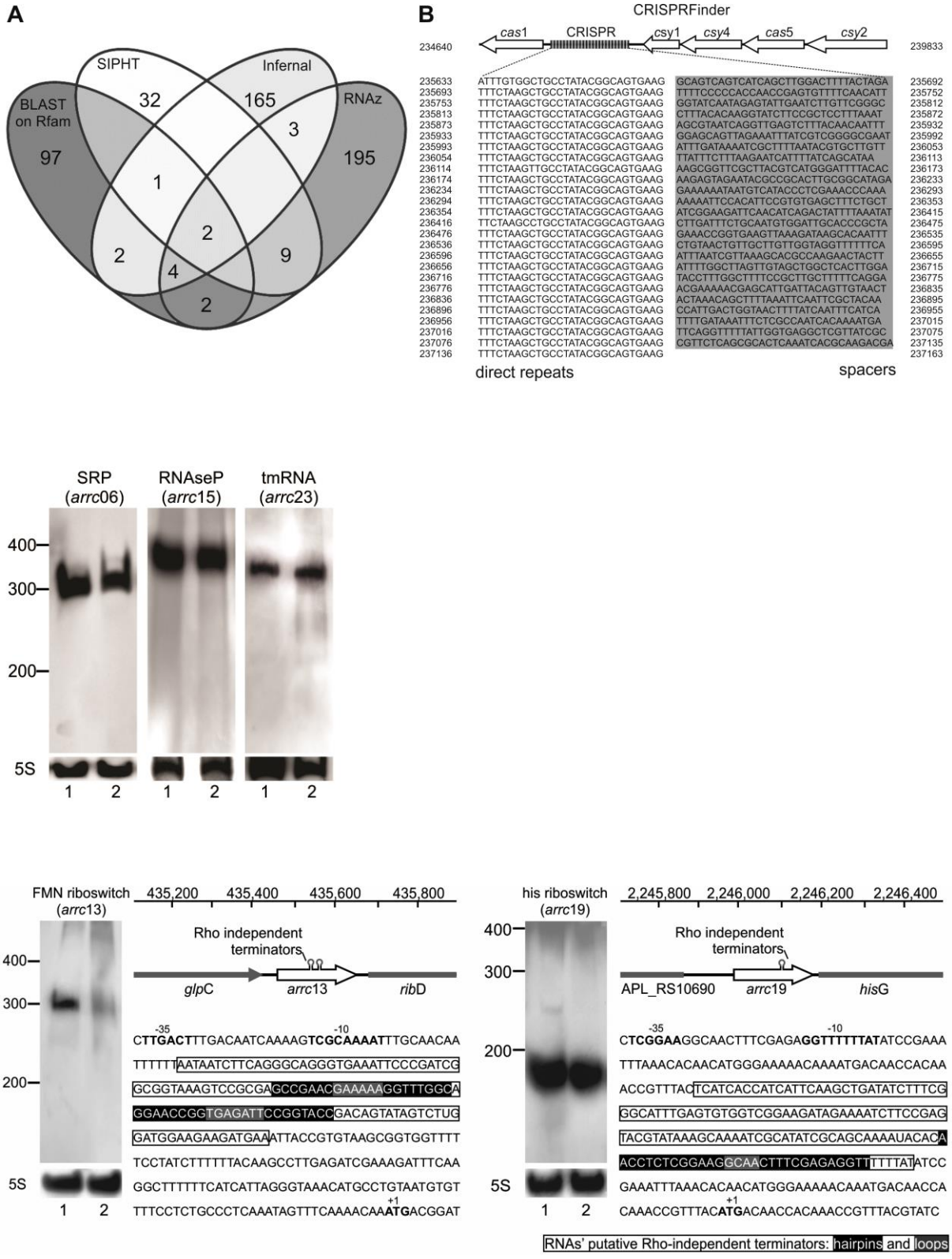
FIGURE 3. Housekeeping regulatory RNAs predicted by bioinformatics for *A. pleuropneumoniae* L20. Expression was validated by Northern blotting under aerobic (1) and anaerobic (2) conditions.

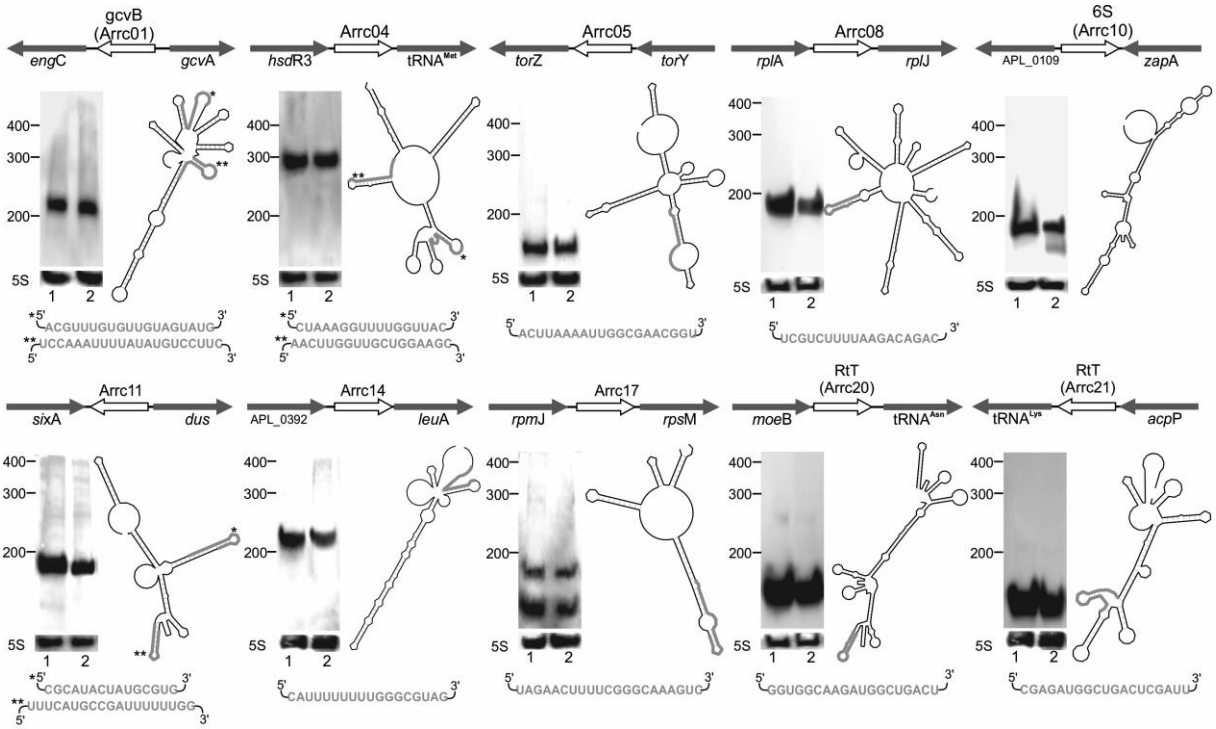
FIGURE 4. *Cis*-acting regulatory RNAs predicted for *A. pleuropneumoniae*. Expression was validated by Northern blotting under aerobic (1) and anaerobic (2) conditions. The genomic context, promoter region, putative terminators, and controlled gene translation starting point are shown for *A*. *pleuropneumoniae* L20.

FIGURE 5. Putative *trans*-acting sRNAs predicted for *A. pleuropneumoniae*. Expression was validated by Northern blotting under aerobic (1) and anaerobic (2) conditions. The predicted secondary structures show the formation of several hairpin regions, defined by the presence of palindromic sequences. Target prediction with TargetRNA2 reveals preferable binding sites (except for RNA polymerase-interacting 6S - Arrc10 sRNA), highlighted in grey. Genomic context is shown for the *A. pleuropneumoniae* reference strain L20.

FIGURE 6. Regulatory network formed by *trans*-acting sRNAs and cognate mRNAs in *A. pleuropneumoniae.* The novel sRNAs described in this work have the potential to bind several mRNA targets, many of which are shared by more than one sRNA, possibly configuring an entangled network of gene regulation in *A. pleuropneumoniae*. The sRNA candidates' names are depicted in white rectangles (Arrcs) and their mRNA connections are illustrated by thick black lines. The other putatively exclusive gene targets are linked to their respective regulatory RNA by gray lines. Targets that may be involved with either virulence or stress resistance are highlighted in black and grey, respectively.


App L20 genome
RNAz
Infernal
BLAST / Rfam
SIPHT
RNAz predictions
Infernal predictions
BLAST / Rfam predictions
SIPHT predictions
BLAST
sequences predicted by at least 2 tools
RNA candidates
primer design
RT-PCR
validated trans-acting sRNAs
TargetRNA2
Gene regulatory network
Northern blotting
sRNA expression and relative abundance
App serotype references (n=11)
App clinical isolates (n=6)
Pasteurellaceae (n=33)
BLAST
sRNA taxonomical dispersion


A
SIPHT
Infernal
RNAz
BLAST on Rfam
32
165
3
97
195
1
2
2
9
4
2
B
CRISPRFinder
cas1
CRISPR
csy1
csy4
cas5
csy2
234640
239833
direct repeats
spacers
SRP (arrc06)
RNAseP (arrc15)
tmRNA (arrc23)
400
300
200
5S
1 2 1 2 1 2
FMN riboswitch (arrc13)
435,200
435,400
435,600
435,800
Rho independent terminators
glpC
arrc13
ribD
his riboswitch (arrc19)
2,245,800
2,246,000
2,246,200
2,246,400
APL_RS10690
arrc19
hisG
RNAs' putative Rho-independent terminators: hairpins and loops


gcvB
(Arrc01)
engC
gcvA
Arrc04
hsdR3
tRNA^Met
Arrc05
torZ
torY
Arrc08
rplA
rplJ
6S
(Arrc10)
APL_0109
zapA
400
300
200
5S
1
2
*5' ACGUUUGUGUUGUAGUAUG 3'
**UCCAAAUUUUAUAUGUCCUUC
5'
3'
*5' CUAAAGGUUUUGGUUAC 3'
**AACUUGGUUGCUGGAAGC
5'
3'
5' ACUUAAAAUUGGCGAACGGU 3'
5' UCGUCUUUUAAGACAGAC 3'
Arrc11
sixA
dus
Arrc14
APL_0392
leuA
Arrc17
rpmJ
rpsM
RtT
(Arrc20)
moeB
tRNA^Asn
RtT
(Arrc21)
tRNA^Lys
acpP
*5' CGCAUACUAUGCGUG 3'
**UUUCAUGCCGAUUUUUUGG
5'
3'
5' CAUUUUUUUUGGGCGUAG 3'
5' UAGAACUUUUCGGGCAAAGUG 3'
5' GGUGGCAAGAUGGCUGACU 3'
5' CGAGAUGGCUGACUCGAUU 3'

SUPPLEMENTAL MATERIAL

Table S1 Oligonucleotides designed for *Actinobacillus pleuropneumoniae* regulatory RNA candidates used in this work.

sRNA target	Primer name	5'→3' sequence	Amplicon (bp)
Arrc01	ARRC01F	TGTTGTGTTTGCATATTGGTCTAGG	122
	ARRC01R	TGGACGGTTATAAACCAAAAAGGT	
Arrc02	ARRC02F	TGCTGATTTCAAGGTAAAAGCG	130
	ARRC02R	GGCTTAAAAGACGAGGGCGA	
Arrc03	ARRC03F	AGGAAAGGCGTATTTGCCGA	119
	ARRC03R	GAACCGACCCTAGCAGTAGC	
Arrc04	ARRC04F	CGCAAAAAGTGCTTGCATTGG	140
	ARRC04R	GCCTTAAAACTGGTTGCGGG	
Arrc05	ARRC05F	CGGTGTGTAAGCGGTCTGAT	103
	ARRC05R	GGATACCGAGCTTGTATGCCT	
Arrc06	ARRC06F	ATGGGGCGTTATTGGTTCCT	105
	ARRC06R	CGTTACCAGCAACCCTCGG	
Arrc07	ARRC07F	AGGTAGCTGGAGAAGAGCGA	182
	ARRC07R	TTCTCCCCTGTCCTTTTGCC	
Arrc08	ARRC08F	AGAGCAAGCTGATGGTGCTT	160
	ARRC08R	CGCTTGCATCGCAAGTAGC	
Arrc09	ARRC09F	AACCGCTTGTCATGAACTCC	121
	ARRC09R	AGAAAAGAGGGGTTAGGGGA	
Arrc10	ARRC10F	CCGTTACTTGTGGTGGTCCT	153
	ARRC10R	TGCGTTACTCGTTACGTCCC	
Arrc11	ARRC11F	TGTCCAATAAATAGGCTTCCCA	126
	ARRC11R	AACTATCCAAATAAAAAGTACGGCT	
Arrc12	ARRC12F	CTACAGGCACATTTCGCAGC	115
	ARRC12R	CGCTTATCGCTAACCGTCTT	
Arrc13	ARRC13F	ATAATCTTCAGGGCAGGGTGA	100
	ARRC13R	ACTGTCGGTACCGGAATCTC	
Arrc14	ARRC14F	ACGACTATCTCTTCGACTGCT	103
	ARRC14R	GCATCAATGTGCGGGCAAAG	
Arrc15	ARRC15F	AGGAACTCAATGGATGGCCC	106
	ARRC15R	TCGATAAGCCGAGTTCTGTCG	
Arrc16	ARRC16F	ACGGGATACATTGGAATTGATAAGG	100
	ARRC16R	TAGGTAATCACTCCAACTCTTACGC	
Arrc17	ARRC17F	TTCTTTCTTGCAAAGAACCCGC	100
	ARRC17R	ATGCTGATCTTGAAAAGCCCG	
Arrc18	ARRC18F	ACGATGAGTCGCAAATTCCC	124
	ARRC18R	AAAGAGAAACTCCGCACAACA	
Arrc19	ARRC19F	TTCGGGCATTTGAGTGTGGT	101
	ARRC19R	AGTTGCCTTCCGAGAGGTTG	
Arrc20	ARRC20F	GCATTTGACGCTAAAACGGT	128
	ARRC20R	AATTAGTGGCTCCTCCTGCG	
Arrc21	ARRC21F	GACCCTTTAGAAGGCGTTGC	115
	ARRC21R	CGCAACGTTAAGGGTCGTTAG	
Arrc22	ARRC22F	GACTCCGAGCTTGTGAACCT	176
	ARRC22R	TGGATTGCATTGGACACCTT	
Arrc23	ARRC23F	TGGATTCGACGGGATTAGCG	179
	ARRC23R	TGGGTGACTTATCGTTGCCC	
rRNA 5S	APP5SF	GCGATGCCCTACTCTCACAT	100
	APP5SR	GAGTGCTGTGGCTCTACCTG	

Table S2 Potential mRNA targets of the putative *Actinobacillus pleuropneumoniae* novel *trans*-acting RNAs described in this work. Putative binding positions in target mRNA 5' UTR were determined using TargetRNA2 (Kery et al. 2014).

RNA ID	Rank	Target gene	Locus	Energy (kcal/mol)	*p*-value	Putative binding position in target mRNA*	Description
Arrc01	1	*arcD*	APL_1082	-18.62	0.000	-29 to -14	arginine/ornithine antiporter
	2	*ilvC*	APL_1853	-18.15	0.000	-62 to -44	ketol-acid reductoisomerase
	3	*ilvI*	APL_0727	-15.74	0.000	-41 to -26	acetolactate synthase 3 catalytic subunit
	4	*dapE*	APL_1873	-14.6	0.001	+3 to 19	succinyl-diaminopimelate desuccinylase
	5	*thrC*	APL_1499	-13.03	0.004	+5 to 20	threonine synthase
	6	*fabG*	APL_1992	-12.73	0.005	-67 to -53	3-ketoacyl-ACP reductase
	7	*ilvD*	APL_0097	-12.55	0.005	-32 to -20	dihydroxy-acid dehydratase
	8	*menA*	APL_1461	-12.1	0.007	-19 to -5	1,4-dihydroxy-2-naphthoate octaprenyltransferase
	9	***hisD***	APL_2020	-11.53	0.010	-9 to 9	histidinol dehydrogenase
	10	*serC*	APL_0702	-11.42	0.011	-10 to +6	phosphoserine aminotransferase
	11	*ftsK*	APL_0618	-11.34	0.011	-48 to -37	DNA translocase FtsK
	12	*dxs*	APL_0207	-11.04	0.013	-5 to +12	1-deoxy-D-xylulose-5-phosphate synthase
	13	*plpA*	APL_0910	-10.77	0.016	-32 to -15	outer membrane lipoprotein 1
	14	*ureA*	APL_1618	-10.64	0.017	-28 to -15	urease subunit gamma
	15	*sfsA*	APL_1737	-10.55	0.018	-8 to +7	sugar fermentation stimulation protein A
	16	***hisC***	APL_2021	-10.32	0.020	2 to 20	histidinol-phosphate aminotransferase
	17	*serB*	APL_1230	-9.95	0.024	-2 to +15	phosphoserine phosphatase
	18	*dapA*	APL_0899	-9.72	0.027	-21 to -2	dihydrodipicolinate synthase
	19	***hisI***	APL_2028	-9.65	0.028	-9 to 12	bifunctional phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP pyrophosphatase
	20	*gcp*	APL_1120	-9.49	0.030	-10 to +9	DNA-binding/iron metalloprotein/AP endonuclease
	21	*hisG*	APL_2019	-9.21	0.034	-25 to -10	ATP phosphoribosyltransferase
	22	*leuC*	APL_0139	-9.14	0.035	-46 to -26	isopropylmalate isomerase large subunit
	23	*gntR*	APL_1667	-9.02	0.037	-3 to +15	HTH-type transcriptional regulator
	24	*secE*	APL_1716	-8.87	0.039	-79 to -64	preprotein translocase subunit SecE
	25	*lrp*	APL_0617	-8.66	0.043	-76 to -63	leucine-responsive transcriptional regulator
	26	*ilvE*	APL_0072	-8.5	0.046	-12 to +8	branched-chain amino acid aminotransferase
	27	*uxaC*	APL_1020	-8.49	0.046	-43 to -30	uronate isomerase
	28	*serA*	APL_1452	-8.4	0.048	-24 to -10	D-3-phosphoglycerate dehydrogenase
	29	*tehB*	APL_1350	-8.35	0.049	-8 to +10	tellurite resistance protein TehB
	30	***trpD***	APL_1165	-8.35	0.049	-69 to -54	anthranilate phosphoribosyltransferase
Arrc02	1	*ilvH*	APL_0728	-15.37	0.000	-15 to -2	acetolactate synthase 3 regulatory subunit
	2	*ptnD*	APL_1393	-14.49	0.001	-18 to -3	PTS system mannose-specific transporter subunit IID
	3	*plsB*	APL_1107	-13.65	0.002	-80 to -64	glycerol-3-phosphate acyltransferase
	4	-	APL_1273	-13.22	0.003	+5 to +19	fimbrial biogenesis and twitching motility protein PilF-like protein
	5	***murI***	APL_1841	-12.8	0.004	-62 to -48	glutamate racemase
	6	*rumB*	APL_1112	-12.45	0.006	+6 to +19	23S rRNA methyluridine methyltransferase
	7	*gmhA*	APL_1364	-12.21	0.007	-76 to -67	phosphoheptose isomerase
	8	***hisD***	APL_2020	-11.19	0.012	-65 to -63	histidinol dehydrogenase
	9	***rpmE***	APL_0982	-10.99	0.014	-51 to -39	50S ribosomal protein L31
	10	*trpA*	APL_0470	-10.88	0.015	-80 to -92	tryptophan synthase subunit alpha
	11	*kpsF*	APL_1576	-10.73	0.016	-15 to -4	arabinose-5-phosphate isomerase
	12	*glpC*	APL_0381	-10.6	0.017		sn-glycerol-3-phosphate dehydrogenase subunit C
	13	*thiQ*	APL_1320	-10.47	0.018	+5 to +19	thiamine transport ATP-binding protein ThiQ
	14	*crr*	APL_1324	-10.41	0.019	-15 to +1	PTS system glucose-specific transporter
	15	*eno*	APL_1113	-10.39	0.019	+1 to +16	phosphopyruvate hydratase
	16	*gnd*	APL_1305	-10.32	0.020	-15 to -1	6-phosphogluconate dehydrogenase
	17	*rpoZ*	APL_1826	-10.18	0.021	-49 to -36	DNA-directed RNA polymerase subunit omega
	18	*napC*	APL_1425	-10.12	0.022	-10 to +5	cytochrome c-type protein NapC
	19	*rdgC*	APL_0161	-9.87	0.025	-45 to -30	recombination associated protein
	20	*ompP4*	APL_0389	-9.86	0.025	-11 to -1	lipoprotein E
	21	*recQ*	APL_1116	-9.62	0.028	-11 to +1	ATP-dependent DNA helicase RecQ
	22	*rnfA*	APL_0165	-9.6	0.028	-67 to -49	Na(+)-translocating NADH-quinone reductase subunit E
	23	*prsA*	APL_0775	-9.39	0.031	-69 to -52	ribose-phosphate pyrophosphokinase
	24	*pheA*	APL_1033	-9.38	0.031	-68 to -51	P-protein

*Position of target mRNA interaction with sRNA given in relation to +1 translation starting point, in accordance to the transcriptome of *Actinobacillus pleuropneumoniae* reference strain L20 (Genbank access NC_009053), considering the region around the 5'UTR of each mRNA. Target genes depicted in bold are located within operons and thus the putative region of interaction between the sRNA and the mRNA is in the interior of the polycistronic RNA.

Table S2 (cont.) Potential mRNA targets of the putative *Actinobacillus pleuropneumoniae* novel *trans*-acting RNAs described in this work. Putative binding positions in target mRNA 5' UTR were determined using TargetRNA2 (Kery et al. 2014).

RNA ID	Rank	Target gene	Locus	Energy (kcal/mol)	*p*-value	Putative binding position in target mRNA*	Description
	26	*accB*	APL_1864	-9.24	0.034	-77 to -68	biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP)
	28	*arcD*	APL_1083	-9.11	0.036	-55 to -39	arginine/ornithine antiporter
	30	*cpdB*	APL_0646	-8.86	0.040	-60 to -48	bifunctional 2',3'-cyclic nucleotide 2'-phosphodiesterase/3'-nucleotidase periplasmic protein
	31	*udp*	APL_1839	-8.78	0.041	+2 to +15	uridine phosphorylase
	32	*csrA*	APL_0653	-8.44	0.047	-11 to +3	carbon storage regulator
	33	*scrK*	APL_2034	-8.35	0.049	+4 to +17	aminoimidazole riboside kinase
	34	*luxS*	APL_1216	-8.35	0.049	-62 to -44	S-ribosylhomocysteinase
	35	*fabG*	APL_1992	-8.34	0.049	-38 to -24	3-ketoacyl-ACP reductase
	37	*oapB*	APL_1404	-8.33	0.049	-17 to -6	opacity associated protein B
Arrc04	1	*utp*	APL_1619	-13.4	0.003	-75 to -62	urea transport protein ApUT
	2	*miaA*	APL_1960	-10.92	0.014	+7 to +20	tRNA delta(2)-isopentenylpyrophosphate transferase
	3	*fabZ*	APL_0408	-10.84	0.015	-17 to -4	(3R)-hydroxymyristoyl-ACP dehydratase
	4	*valS*	APL_1502	-10.74	0.016	+5 to +14	valyl-tRNA synthetase
	5	*fumC*	APL_1757	-10.65	0.017	-69 to -57	fumarate hydratase
	6	***malE***	APL_1237	-10.2	0.021	-26 to -14	maltose ABC transporter periplasmic protein
	7	*ccp*	APL_1379	-10.09	0.022	-68 to -58	cytochrome c peroxidase
	8	*cysJ*	APL_1843	-9.85	0.025	-2 to +13	sulfite reductase [NADPH] flavoprotein alpha-component (SIR-FP)
	9	*cysB*	APL_0133	-9.84	0.025	-62 to -54	transcriptional regulator CysB
	10	*mioC*	APL_1563	-9.82	0.026	-13 to +2	flavodoxin
	11	*rimK*	APL_0484	-9.8	0.026	-76 to -62	ribosomal protein S6 modification protein
	12	*gnd*	APL_1305	-9.77	0.026	-23 to -8	6-phosphogluconate dehydrogenase
	13	***moaC***	APL_0691	-9.64	0.028	+1 to +15	molybdenum cofactor biosynthesis protein C
	14	*crr*	APL_1324	-9.63	0.028	-11 to +2	PTS system glucose-specific transporter
	15	*fadI*	APL_0887	-9.47	0.030	-39 to -25	3-ketoacyl-CoA thiolase
	16	*pheT*	APL_0609	-9.25	0.033	-12 to +7	phenylalanyl-tRNA synthetase subunit beta
	17	*kpsF*	APL_1576	-9	0.037	-14 to +6	arabinose-5-phosphate isomerase
	18	*nusA*	APL_0638	-8.81	0.040	+8 to +18	transcription elongation factor NusA
	19	*mipB*	APL_0062	-8.8	0.041	-3 to +8	transaldolase B
	20	*rimO*	APL_1636	-8.78	0.041	-3 to +12	30S ribosomal protein S12 methylthiotransferase
	21	*tyrR*	APL_0797	-8.67	0.043	-69 to -53	transcriptional regulatory protein TyrR
	22	*argD*	APL_0244	-8.67	0.043	+1 to +19	acetylornithine aminotransferase
	23	*bioA*	APL_0942	-8.62	0.044	+1 to +13	adenosylmethionine-8-amino-7-oxononanoate aminotransferase
	24	*ulaD*	APL_1698	-8.5	0.046	+3 to +13	3-keto-L-gulonate-6-phosphate decarboxylase
	25	*nrfB*	APL_0101	-8.49	0.046	-63 to -51	cytochrome c nitrite reductase pentaheme subunit
Arrc05	1	*glmS*	APL_1631	-17.08	0.000	-78 to -63	glucosamine--fructose-6-phosphate aminotransferase
	2	*trkH*	APL_1590	-15.99	0.000	-54 to -35	Trk system potassium uptake protein TrkH
	3	*rnb*	APL_0757	-15.07	0.001	-80 to -67	exoribonuclease II
	4	*recJ*	APL_0459	-13	0.004	-32 to -14	single-stranded-DNA-specific exonuclease RecJ
	5	*arcD*	APL_1082	-12.74	0.005	-44 to -30	arginine/ornithine antiporter
	6	*accC*	APL_1865	-12.53	0.005	-63 to -47	acetyl-CoA carboxylase biotin carboxylase subunit
	7	*dut*	APL_1968	-11.22	0.012	-27 to -15	deoxyuridine 5'-triphosphate nucleotidohydrolase
	8	*dnaG*	APL_1474	-10.38	0.019	-32 to -17	DNA primase
	9	*rnpA*	APL_1939	-10.31	0.020	-80 to -66	ribonuclease P
	10	*xseA*	APL_0817	-10.06	0.023	-16 to -4	exodeoxyribonuclease VII large subunit
	11	*gpt*	APL_0255	-10.03	0.023	-19 to -3	xanthine-guanine phosphoribosyltransferase
	12	*sbcB*	APL_0673	-9.99	0.023	-56 to -40	exonuclease I
	13	*hemE*	APL_0112	-9.87	0.025	-45 to -30	uroporphyrinogen decarboxylase

*Position of target mRNA interaction with sRNA given in relation to +1 translation starting point, in accordance to the transcriptome of *Actinobacillus pleuropneumoniae* reference strain L20 (Genbank access NC_009053), considering the region around the 5'UTR of each mRNA. Target genes depicted in bold are located within operons and thus the putative region of interaction between the sRNA and the mRNA is in the interior of the polycistronic RNA.

Table S2 (cont.) Potential mRNA targets of the putative *Actinobacillus pleuropneumoniae* novel *trans*-acting RNAs described in this work. Putative binding positions in target mRNA 5' UTR were determined using TargetRNA2 (Kery et al. 2014).

RNA ID	Rank	Target gene	Locus	Energy (kcal/mol)	*p*-value	Putative binding position in target mRNA*	Description
	14	*cca*	APL_0915	-9.66	0.028	-63 to -54	multifunctional tRNA nucleotidyl transferase/2'3'-cyclic phosphodiesterase/2'nucleotidase/phosphatase
	15	***malQ***	APL_1240	-9.59	0.029	-48 to -38	4-alpha-glucanotransferase
	16	*lyx*	APL_1564	-9.48	0.030	-70 to -61	L-xylulose kinase
	17	*pyrG*	APL_0136	-9.37	0.032	-58 to -47	CTP synthetase
	18	*copA*	APL_1265	-9.19	0.034	-60 to -52	copper-transporting P-type ATPase
	19	-	APL_0271	-8.58	0.044	+1 to +10	chelated iron transport system ATP-binding protein
	20	*uspA*	APL_0655	-8.56	0.045	-29 to -13	universal stress protein A-like protein
	21	*lacZ*	APL_0997	-8.52	0.045	-42 to -32	beta-galactosidase
	22	*ostA*	APL_0962	-8.5	0.046	-55 to -44	organic solvent tolerance protein
	23	*dipZ*	APL_1359	-8.42	0.047	-18 to -1	thiol:disulfide interchange protein
	24	*pheT*	APL_0609	-8.38	0.048	-60 to -51	phenylalanyl-tRNA synthetase subunit beta
Arrc07	1	*napF*	APL_1431	-14.11	0.002	-80 to -68	ferredoxin-type protein NapF
	2	*menD*	APL_1750	-12.32	0.006	+6 to +17	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase
	3	*pgaA*	APL_1921	-11.71	0.009	-80 to -72	biofilm PGA synthesis protein PgaA
	4	*rumB*	APL_1112	-11.33	0.011	-22 to -14	23S rRNA methyluridine methyltransferase
	5	*focA*	APL_1037	-10.1	0.022	+4 to +12	formate transporter
	6	*mtfA*	APL_0685	-10.07	0.023	-26 to -12	RNA 2'-O-ribose methyltransferase
	7	*yedE*	APL_1977	-10.02	0.023	-64 to -50	inner membrane protein
	8	*cspC*	APL_0118	-9.45	0.030	-56 to -48	cold shock-like protein CspC
	9	*vacB*	APL_1478	-9.2	0.034	+7 to +19	ribonuclease R
	10	*pgsA*	APL_0275	-9.18	0.034	-30 to -16	phosphatidylglycerophosphate synthase
	11	*msbA*	APL_0778	-9.16	0.035	+12 to +20	lipid transporter ATP-binding protein/permease
	12	*grpE*	APL_0367	-9.06	0.036	-80 to -70	heat shock protein GrpE
	13	*mreC*	APL_0436	-8.99	0.037	-36 to -28	rod shape-determining protein MreC
	14	*purE*	APL_0659	-8.72	0.042	+11 to +12	phosphoribosylaminoimidazole carboxylase catalytic subunit
	15	*pepB*	APL_0388	-8.55	0.045	-64 to -55	aminopeptidase B
Arrc08	1	*pta*	APL_0644	-11.72	0.009	+1 to +11	phosphate acetyltransferase
	2	*ung*	APL_0362	-10.59	0.017	+9 to +20	uracil-DNA glycosylase
	3	*lpxD*	APL_0409	-10.26	0.021	-42 to -28	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase
	4	*sufI*	APL_1487	-9.39	0.031	+10 to +20	protein SufI
	5	*aroQ*	APL_1862	-9.19	0.034	+6 to +20	3-dehydroquinate dehydratase
	6	*hyaB*	APL_1334	-9.15	0.035	-47 to -35	hydrogenase 2 large subunit
	7	*tmk*	APL_1817	-8.93	0.038	+4 to +16	thymidylate kinase
	8	*ftsY*	APL_1346	-8.43	0.047	+2 to +12	cell division protein FtsY
Arrc11	1	*copA*	APL_1265	-16	0.000	-5 to -14	copper-transporting P-type ATPase
	2	***murG***	APL_0018	-12.89	0.004	-1 to -15	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase
	3	***rpsK***	APL_1782	-11.59	0.010	-1 to +15	30S ribosomal protein S11
	4	*ispH*	APL_1520	-11.22	0.012	-11 to +3	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
	5	*ccmF*	APL_1367	-10.77	0.016	-14 to -5	cytochrome c-type biogenesis protein CcmF
	6	*fdhE*	APL_0896	-9.85	0.025	-77 to -68	formate dehydrogenase accessory protein FdhE
	7	***metN***	APL_0912	-9.83	0.025	+4 to +14	DL-methionine transporter ATP-binding subunit
	8	*ung*	APL_0362	-9.72	0.027	+7 to +17	uracil-DNA glycosylase
	9	*dam*	APL_0192	-9.02	0.037	-80 to -68	DNA adenine methylase
	10	*ftsK*	APL_0618	-8.84	0.040	-80 to -66	DNA translocase FtsK
	11	*ulaG*	APL_1701	-8.65	0.043	-2 to +11	L-ascorbate 6-phosphate lactonase
	12	*djlA*	APL_0306	-8.63	0.043	+2 to +16	Dna-J like membrane chaperone protein
	13	*radA*	APL_0881	-8.54	0.045	-1 to 15	DNA repair protein RadA
	14	*apbE*	APL_0156	-8.52	0.046	-27 to -18	thiamine biosynthesis lipoprotein ApbE
	15	*eno*	APL_1113	-8.48	0.046	-1 to +15	phosphopyruvate hydratase

*Position of target mRNA interaction with sRNA given in relation to +1 translation starting point, in accordance to the transcriptome of *Actinobacillus pleuropneumoniae* reference strain L20 (Genbank access NC_009053), considering the region around the 5'UTR of each mRNA. Target genes depicted in bold are located within operons and thus the putative region of interaction between the sRNA and the mRNA is in the interior of the polycistronic RNA.

Table S2 (cont.) Potential mRNA targets of the putative *Actinobacillus pleuropneumoniae* novel *trans*-acting RNAs described in this work. Putative binding positions in target mRNA 5' UTR were determined using TargetRNA2 (Kery et al. 2014).

RNA ID	Rank	Target gene	Locus	Energy (kcal/mol)	*p*-value	Putative binding position in target mRNA *	Description
Arrc14	1	*proA*	APL_1951	-15.29	0.001	-79 to -63	gamma-glutamyl phosphate reductase
	2	*htpG*	APL_0987	-13.19	0.003	-77 to -63	heat shock protein 90
	3	*hcp*	APL_1546	-12.95	0.004	+5 to +20	hydroxylamine reductase
	4	*ompP2A*	APL_0006	-12.62	0.005	+1 to +17	outer membrane protein P2
	5	*fdx*	APL_0924	-12.12	0.007	+8 to +20	2Fe-2S ferredoxin
	6	*cbiK*	APL_1624	-11.88	0.008	-80 to -69	periplasmic binding protein CbiK
	7	*mutT*	APL_0241	-11.01	0.014	+1 to +16	mutator mutT protein
	8	*grcA*	APL_0361	-10.98	0.014	-70 to -60	autonomous glycyl radical cofactor GrcA
	9	*ispZ*	APL_0972	-10.95	0.014	-80 to -66	intracellular septation protein A
	10	*bioD*	APL_0614	-10.38	0.019	-56 to -44	dithiobiotin synthetase
	11	*znuC*	APL_0456	-10.12	0.022	-54 to -42	high-affinity zinc uptake system ATP-binding protein ZnuC
	12	*glmS*	APL_1631	-10.03	0.023	-1 to +9	glucosamine--fructose-6-phosphate aminotransferase
	13	*galT*	APL_0994	-9.92	0.024	-53 to -40	galactose-1-phosphate uridylyltransferase
	14	*wecB*	APL_1552	-9.84	0.025	-73 to -58	UDP-N-acetylglucosamine 2-epimerase
	15	*rbsD*	APL_1669	-9.82	0.026	+1 to +16	D-ribose pyranase
	16	*frdA*	APL_1529	-9.8	0.026	+2 to +12	fumarate reductase flavoprotein subunit
	17	*tadD*	APL_0549	-9.78	0.026	-64 to -50	tight adherence protein D
	18	*zwf*	APL_1311	-9.77	0.026	-79 to -64	glucose-6-phosphate 1-dehydrogenase
	19	*glmU*	APL_0588	-9.47	0.030	-59 to -41	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase
	20	*nrfX*	APL_1051	-9.37	0.032	+1 to +14	disulfide bound formation protein DsbE
	21	*leuS*	APL_0872	-9.07	0.036	-61 to -50	leucyl-tRNA synthetase
	22	*ompP2*	APL_0649	-9.03	0.037	+1 to +14	Outer membrane protein P2 precursor (OMP P2)
	23	*tadE*	APL_0548	-8.76	0.041	+8 to +20	tight adherence protein E
	24	*torD*	APL_1797	-8.51	0.046	-80 to -66	chaperone protein TorD
	25	*afuB_2*	APL_0564	-8.51	0.046	+9 to +19	ferric transport system permease protein
	26	*mglB*	APL_1420	-8.32	0.049	+1 to +17	D-galactose-binding periplasmic protein
Arrc17	1	***rpsM***	APL_1781	-15.9	0.000	-46 to -27	30S ribosomal protein S13
	2	*arsR*	APL_1090	-14.2	0.001	-80 to -63	transcriptional regulator
	3	*menC*	APL_0353	-12.1	0.007	-79 to -65	O-succinylbenzoate synthase
	4	*rsmB*	APL_1560	-11.75	0.009	-76 to -58	ribosomal RNA small subunit methyltransferase B
	5	*fdhE*	APL_0896	-11.47	0.010	+1 to +14	formate dehydrogenase accessory protein FdhE
	6	*frdB*	APL_1528	-11.3	0.012	-18 to -15	fumarate reductase iron-sulfur subunit
	7	*cysW*	APL_1847	-10.12	0.022	+1 to +15	sulfate transport system permease protein cysW
	8	*dppD*	APL_0067	-9.58	0.029	-23 to -15	dipeptide transporter ATP-binding subunit
	9	*coaE*	APL_0876	-9.49	0.030	-66 to -56	dephospho-CoA kinase
	10	*priA*	APL_1032	-9.28	0.033	-65 to -48	primosome assembly protein PriA
	11	*rnfB*	APL_0166	-9.18	0.034	+1 to +19	electron transport complex protein RnfB
	12	*pbpB*	APL_1823	-9.16	0.035	-20 to -1	penicillin-binding protein 1B (PBP1b)
	13	*purE*	APL_0659	-8.77	0.041	+12 to +20	phosphoribosylaminoimidazole carboxylase catalytic subunit
	14	*rnfA*	APL_0165	-8.7	0.042	-64 to -53	Na(+)-translocating NADH-quinone reductase subunit E
	15	*thrA*	APL_0250	-8.5	0.046	-39 to -26	bifunctional aspartokinase I/homoserine dehydrogenase I
	16	*dksA*	APL_0175	-8.39	0.048	-13 to -5	DnaK suppressor protein
Arrc20	1	*pmbA*	APL_0729	-13.03	0.004	-45 to -34	antibiotic maturation factor
	2	-	APL_1678	-11.76	0.009	-35 to -23	ferredoxin
	3	*corA*	APL_1981	-10.53	0.018	-44 to -30	magnesium/nickel/cobalt transporter CorA
	4	*pta*	APL_0644	-9.64	0.028	-3 to +9	phosphate acetyltransferase
	5	*fldA*	APL_1219	-9.28	0.033	-79 to -66	flavodoxin FldA
	6	*lctP*	APL_0447	-9.14	0.035	+5 to +19	L-lactate permease
	7	*ureA*	APL_1618	-8.47	0.046	-79 to -61	urease subunit gamma
	8	*fumC*	APL_1757	-8.32	0.049	+6 to +17	fumarate hydratase

*Position of target mRNA interaction with sRNA given in relation to +1 translation starting point, in accordance to the transcriptome of *Actinobacillus pleuropneumoniae* reference strain L20 (Genbank access NC_009053), considering the region around the 5'UTR of each mRNA. Target genes depicted in bold are located within operons and thus the putative region of interaction between the sRNA and the mRNA is in the interior of the polycistronic RNA.

Table S2 (cont.) Potential mRNA targets of the putative *Actinobacillus pleuropneumoniae* novel *trans*-acting RNAs described in this work. Putative binding positions in target mRNA 5' UTR were determined using TargetRNA2 (Kery et al. 2014).

RNA ID	Rank	Target gene	Locus	Energy (kcal/mol)	*p*-value	Putative binding position in target mRNA *	Description
Arrc21	1	***ubiC***	APL_1840	-11.02	0.014	-74 to -56	4-hydroxybenzoate synthetase
	2	*apxIIA*	APL_0956	-10.73	0.016	-51 to -43	RTX-II toxin determinant A
	3	*bioA*	APL_0942	-10.26	0.020	+1 to +17	adenosylmethionine-8-amino-7-oxononanoate aminotransferase
	4	*anmK*	APL_1535	-9.57	0.029	+12 to +20	anhydro-N-acetylmuramic acid kinase
	5	*cysK*	APL_0248	-9.18	0.034	-10 to +6	cysteine synthase
	6	*guaA*	APL_0592	-9.08	0.036	-11 to +8	GMP synthase
	7	*malE*	APL_1237	-8.6	0.044	-59 to -50	maltose ABC transporter periplasmic protein
	8	*menA*	APL_1461	-8.49	0.046	-80 to -67	1,4-dihydroxy-2-naphthoate octaprenyltransferase
	9	*coaE*	APL_0876	-8.3	0.050	-6 to +9	dephospho-CoA kinase

*Position of target mRNA interaction with sRNA given in relation to +1 translation starting point, in accordance to the transcriptome of *Actinobacillus pleuropneumoniae* reference strain L20 (Genbank access NC_009053), considering the region around the 5'UTR of each mRNA. Target genes depicted in bold are located within operons and thus the putative region of interaction between the sRNA and the mRNA is in the interior of the polycistronic RNA.

Table S3 Distribution of the small RNA candidates described in this work among *Actinobacillus pleuropneumoniae* strains and clinical isolates and other *Pasteurellaceae.*

Species	RNAs																						
	Arrc01	**Arrc02**	**Arrc03**	**Arrc04**	**Arrc05**	**Arrc06**	**Arrc07**	**Arrc08**	**Arrc09**	**Arrc10**	**Arrc11**	**Arrc12**	**Arrc13**	**Arrc14**	**Arrc15**	**Arrc16**	**Arrc17**	**Arrc18**	**Arrc19**	**Arrc20**	**Arrc21**	**Arrc22**	**Arrc23**
Actinobacillus pleuropneumoniae																							
Shope4074	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
S1536	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
JL03	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
M62	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
L20	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Femo	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
AP76	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
CVJ13261	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
D13039	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
56153	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1096	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
N273	x	-	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
518	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
5651	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
597	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
780	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
460	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1022	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Actinobacillus succinogenes																							
130Z	-	-	-	-	-	-	-	-	-	-	-	x	-	-	x	-	x	-	-	-	x	-	x
Actinobacillus suis																							
H91-0380	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Aggregatibacter actinomycetemcomitans																							
ANH9381	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x
D11S-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x
D7S-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x
Aggregatibacter aphrophilus																							
NJ8700	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x
Gallibacterium anatis																							
UMN179	-	-	-	-	-	-	-	-	-	-	-	x	-	-	x	-	x	-	-	-	x	-	x
Haemophilus ducrey																							
35000HP	-	x	-	-	-	-	x	x	-	-	-	x	x	-	x	x	x	-	-	-	x	-	x
Haemophilus influenzae																							
86-028NP	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
10810	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
F3031	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
F3047	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
PittEE	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
PittGG	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
R2864	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
R2866	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
RdKW20	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	-	-	x	-	x
Haemophilus parainfluenzae																							
T3T1	-	-	-	-	-	-	x	-	-	-	-	-	-	-	x	-	-	-	-	-	-	-	x

Table S3 (cont.) Distribution of the small RNA candidates described in this work among *Actinobacillus pleuropneumoniae* strains and clinical isolates and other *Pasteurellaceae*.*

	RNAs																						
Species	**Arrc01**	**Arrc02**	**Arrc03**	**Arrc04**	**Arrc05**	**Arrc06**	**Arrc07**	**Arrc08**	**Arrc09**	**Arrc10**	**Arrc11**	**Arrc12**	**Arrc13**	**Arrc14**	**Arrc15**	**Arrc16**	**Arrc17**	**Arrc18**	**Arrc19**	**Arrc20**	**Arrc21**	**Arrc22**	**Arrc23**
Haemophilus parasuis																							
SH0165	x	-	-	-	-	-	x	-	-	-	-	-	x	-	x	-	x	-	-	-	x	-	x
Haemophilus somnus																							
129T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	x	-	-	-	x	-	x
2336	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	x	-	-	-	x	-	x
Mannheimia haemolytica																							
M42548	-	-	x	-	-	-	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
D153	-	-	x	-	-	-	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
D171	-	-	x	-	-	-	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
D174	-	-	x	-	-	-	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
USDA-ARS-USMARC-183	-	-	x	-	-	-	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
Mannheimia succiniciproducens																							
MBEL55E	-	-	x	-	-	x	-	-	-	-	-	x	-	-	x	-	-	-	-	-	x	-	x
Mannheimia varigena																							
USDA-ARS-USMARC-1261	-	-	-	-	-	x	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
USDA-ARS-USMARC-1296	-	-	x	-	-	x	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
USDA-ARS-USMARC-1312	-	-	x	-	-	x	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
USDA-ARS-USMARC-1388	-	-	x	-	-	-	x	x	-	x	-	-	x	-	x	x	x	-	-	x	x	-	x
Pasteurella multocida																							
Pm70	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x
HN06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-	-	x	-	x

* "x" denotes the presence of a determined RNA, while "-" denotes the absence.

Table S4. Genbank accession numbers of the *Pasteurellaceae* genomes analyzed in this work.

Species	**Strain (Genbank accession number)**
Actinobacillus pleuropneumoniae	Reference strains: Shope4074, serotype 1 (ADOD00000000); S1536, serotype 2 (ADOE00000000); JL03, serotype 3 (NC_010278); M62, serotype 4 (ADOF00000000); L20, serotype 5 (NC_009053); Femo, serotype 6 (ADOG00000000); AP76, serotype 7 (NC_010939); CVJ13261, serotype 9 (ADOI00000000); D13039, serotype 10 (ADOJ00000000); 56153, serotype 11 (ADOK00000000); 1096, serotype 12 (ADOL00000000); N273, serotype 13 (ADOM00000000).
	Clinical isolates: 518 (JSVZ00000000), 5651 (JSVY00000000), 597 (JSVX00000000), 780 (JSVV00000000), 460 (JSVG00000000), 1022 (JSVF00000000), all serotype 8.
Actinobacillus succinogenes	130Z (NC_009655)
Actinobacillus suis	H91-0380 (NC_018690)
Aggregatibacter actinomycetemcomitans	ANH9381 (CP003099), D11S-1 (CP001733), D7S-1 (CP003496)
Aggregatibacter aphrophilus	NJ8700 (CP001607)
Gallibacterium anatis	UMN179 (NC_015460)
Haemophilus ducreyi	35000HP (NC_002940)
Haemophilus influenzae	RdKW20 (NC_000907), 10810 (NC_016809), F3047 (NC_014922), 86-028NP (NC_007146), F3031 (NC_014920), PittGG (CP000672), PittEE (CP000671), R2846 (CP002276), R2866 (CP002277)
Haemophilus parainfluenzae	T3T1 (NC_015964)
Haemophilus parasuis	SH0165 (NC_011852)
Haemophilus somnus	129T (NC_008309), 2336 (NC_010519)
Mannheimia haemolytica	M42548 (NC_021082), D153 (NC_021743), D171 (NC_021738), D174 (NC_021739), USDA-ARS-USMARC-183 (NC_020833)
Mannheimia succiniciproducens	MBEL55E (NC_006300)
Mannheimia varigena	USDA-ARS-USMARC-1261 (CP006942); USDA-ARS-USMARC-1296 (CP006943); USDA-ARS-USMARC-1312 (CP006944); USDA-ARS-USMARC-1388 (CP006953)
Pasteurella multocida	Pm70 (NC_002663), HN06 (NC_017027)

A) FMN Riboswitch – Rfam Accession RF00050

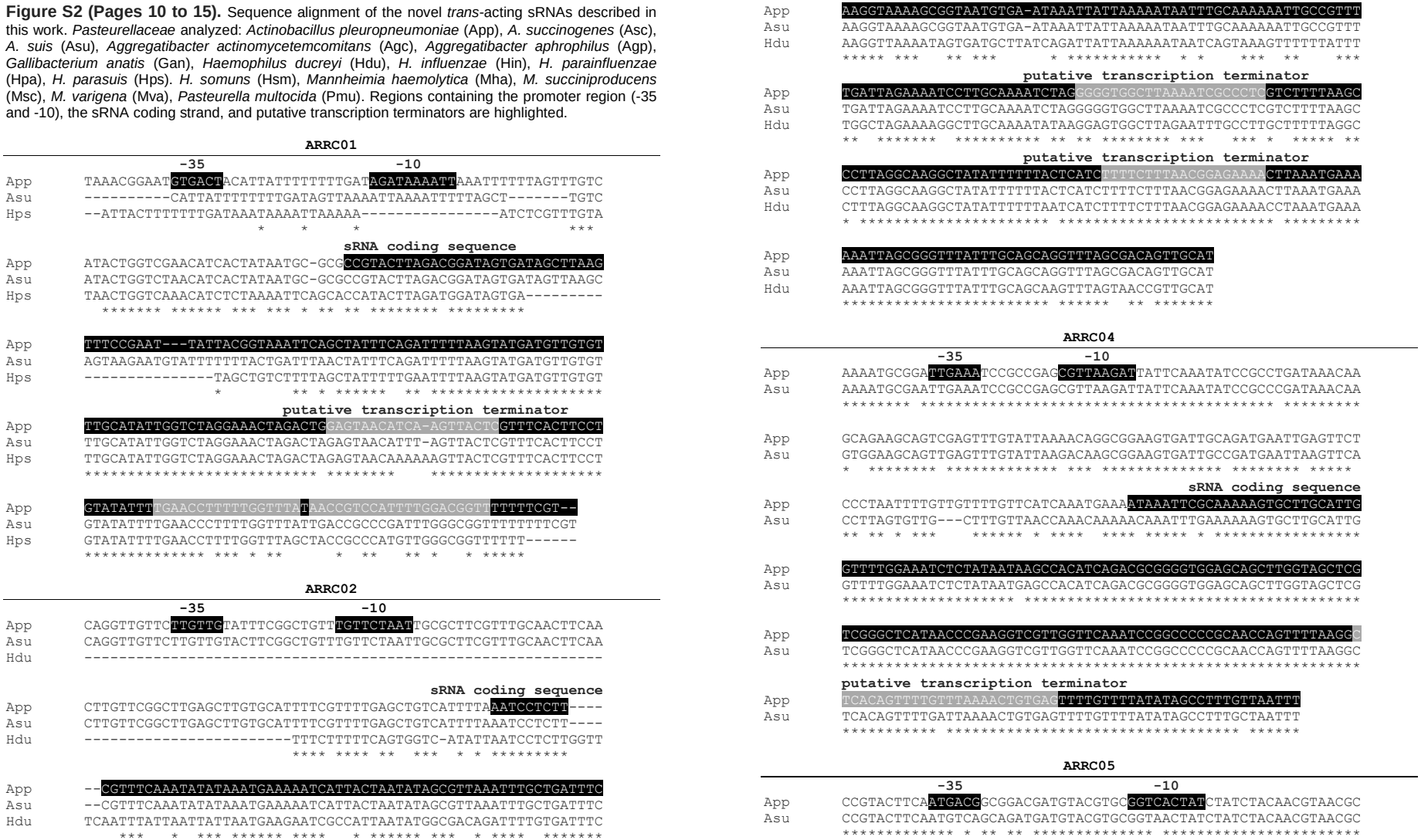
```
SS.           ::(((((((((,,,,,,<<<_______>>>,,,,,<<<<_____________________>>>
FMN-Rfam      AAUCGUCCUCAGGGCAGGGUGAAAAUUCCCUACCGGCGGUAAUUAAUUAAAUUUAAGCCCG
FMN-App       AAUAAUCUUCAGGGCAGGGUGAAAAUUCCCGAUCGGCGGUAA-------------AGUCCG
              ***..** ********************* * *********             ** ***
SS.           >,,,,,,,,,,,,,,,,,,,,,,,,,,,,<<<<<_______>>>>>,,,,,<<<_____
FMN-Rfam      CGAGCGUUUAAAAUUUUAAAGUCAGCAGAUCCGGUGAAAUUCCGGAGCCGACGGUAUAGU
FMN-App       CGAGCCGAACGA-AAAAGGUUUGGCAGGAACCGGUGAGAUUCCGGUACCGACAGUAUAGU
              *****    ..*     ..  * . ..** *******.******* .*****.*******
SS.           >>>,,,,,,,))))))))::
FMN-Rfam      CCGGAUGGAAGAGGACGAAA
FMN-App       CUGGAUGGAAGAAGAUGAAA
              * **********.** ****
```

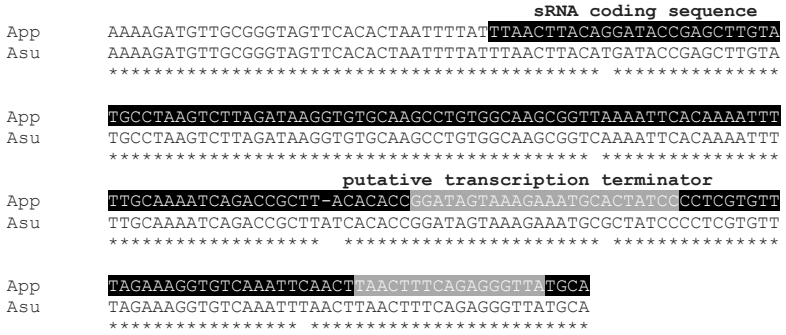
B) His Attenuator – Rfam Accession RF00514

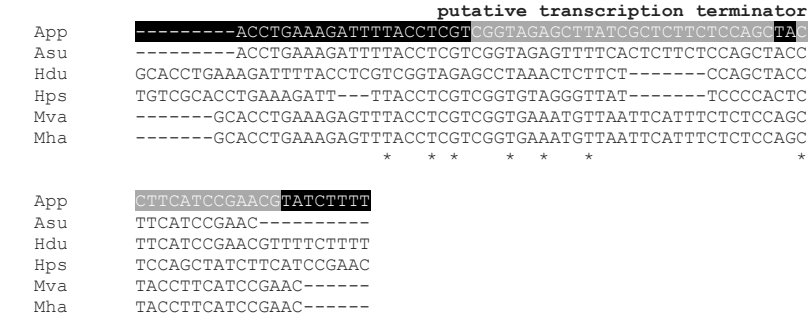
```
SS.           :::::::<<<<<<<.--<<<<<________>>>>>-.>>>>-.->>>.<<<<<<<<___
His-Rfam      UCAUCACCACCAUC-AUCCUGACUAGUCUUUCAGG-AGAUGU-UGUG-CUGGAAGGCAA
His-App       UCAUCACCAUCAUUCAAGCUGA--UAUCUUUCGGGCAUUUGAGUGUGGUCGGAAGAUAG
              ********* ***  *  ****   .******.** *. **  ****   *****  *
SS.           _____>>>>>>>>--.....................,,,,,,,,,,,,,,,,,<<<.<<<
His-Rfam      UUAUACCUUCCAGAG--------------------UGAUAUAAAAACAAAUUAAA-AAA
His-App       AAA-AUCUUCCGAGUACGUAUAAAGCAAAAUCGCA---UAUCGCAGCAAAAUACACAAC
                * * *****...                          *** . *.**** ** * **
SS.           <<<<<<<<<<_____>>>>>>>>>>>>>>>>>:::
His-Rfam      CCCUCGGGAGAUCAUCUCCCGAGGGUUUUUUUUU
His-App       CUCUCGGAAG-GCAACUUUCGAGAGGUUUUUUAU
              * *****.**  ** **  ****.* ****** *
```

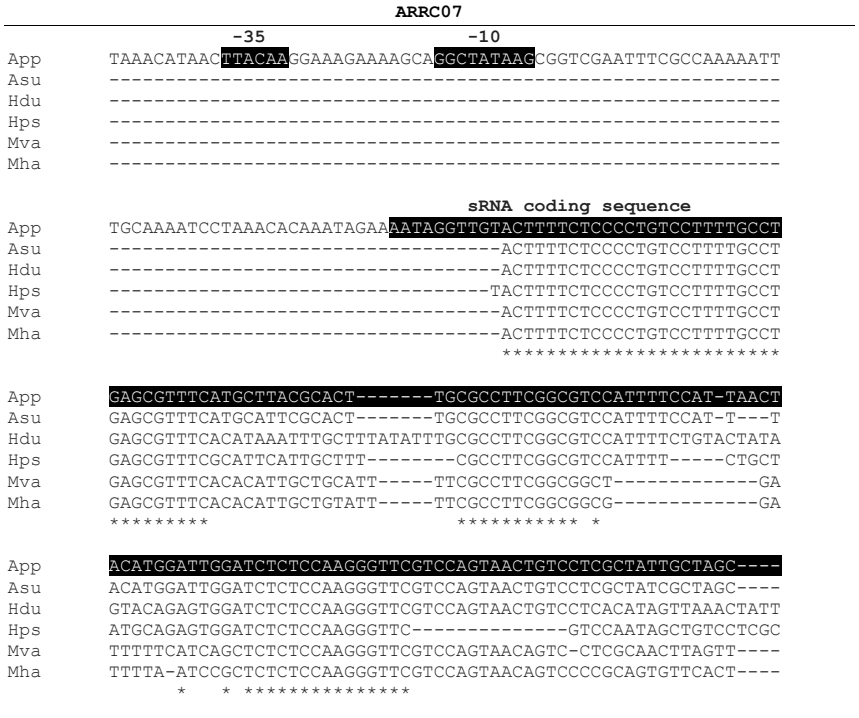
Figure S1. Nucleotide sequence alignments and secondary structures (SS) of the *cis*-acting sRNAs whose expressions were confirmed in this work. Alignments were made between the element predicted for *Actinobacillus pleuropneumoniae* (App) and the consensus sequence available from the Rfam database. Figure shows alignments for both the FMN riboswitch (A) and the Histidine attenuator (B). Sequential " **<** and **>** " represent sequences that pair with each other, separated by loops, represented by "____". Regions not involved with secondary structures are represented by " **,** " when in the interior of the sequence or by " **:** " when in the extremities. Distant regions that pair with each other are represented by " **(** and **)** ". Expression platforms are highlighted in yellow.

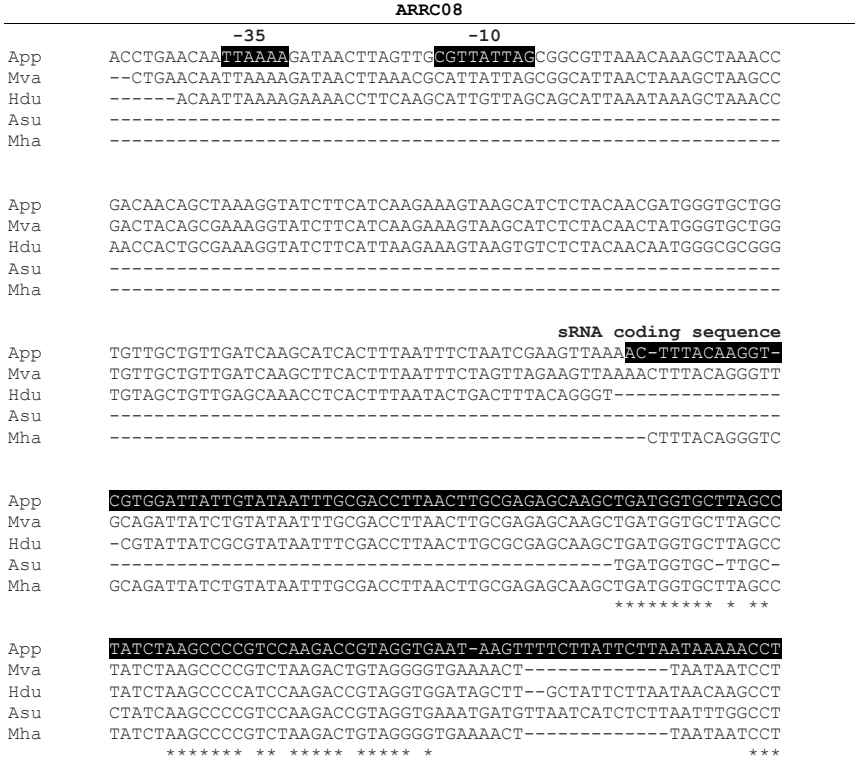
Figure S2 (Pages 10 to 15). Sequence alignment of the novel *trans*-acting sRNAs described in this work. *Pasteurellaceae* analyzed: *Actinobacillus pleuropneumoniae* (App), *A. succinogenes* (Asc), *A. suis* (Asu), *Aggregatibacter actinomycetemcomitans* (Agc), *Aggregatibacter aphrophilus* (Agp), *Gallibacterium anatis* (Gan), *Haemophilus ducreyi* (Hdu), *H. influenzae* (Hin), *H. parainfluenzae* (Hpa), *H. parasuis* (Hps). *H. somuns* (Hsm), *Mannheimia haemolytica* (Mha), *M. succiniproducens* (Msc), *M. varigena* (Mva), *Pasteurella multocida* (Pmu). Regions containing the promoter region (-35 and -10), the sRNA coding strand, and putative transcription terminators are highlighted.

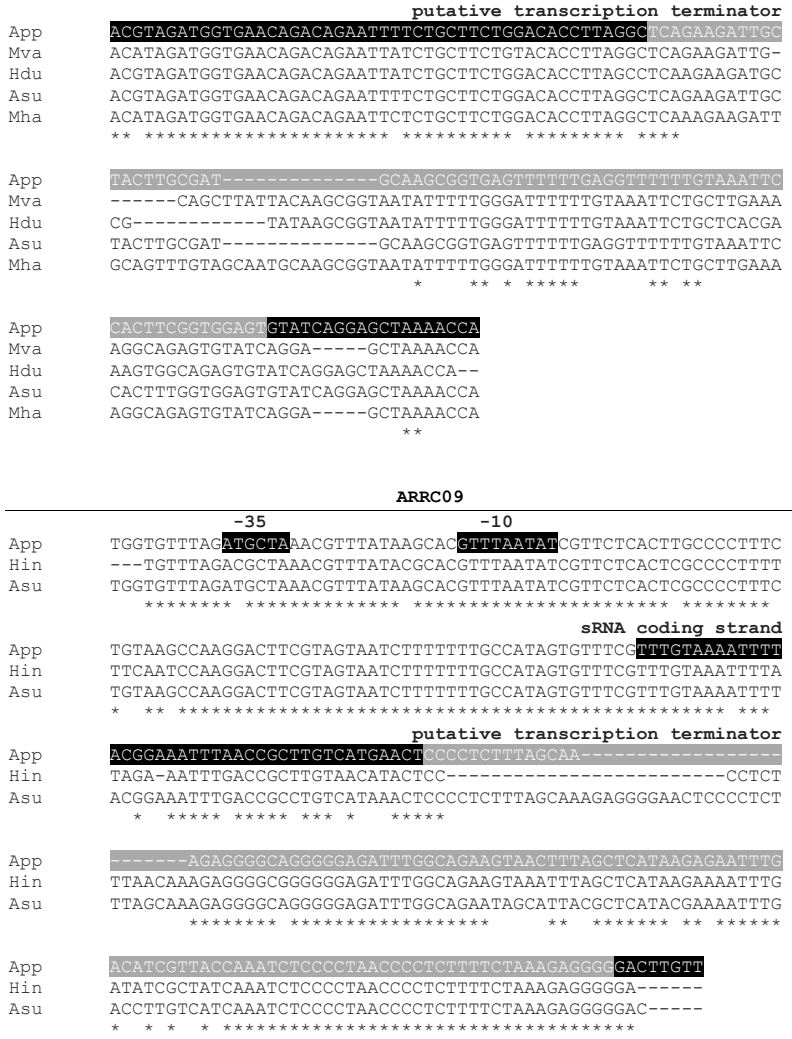


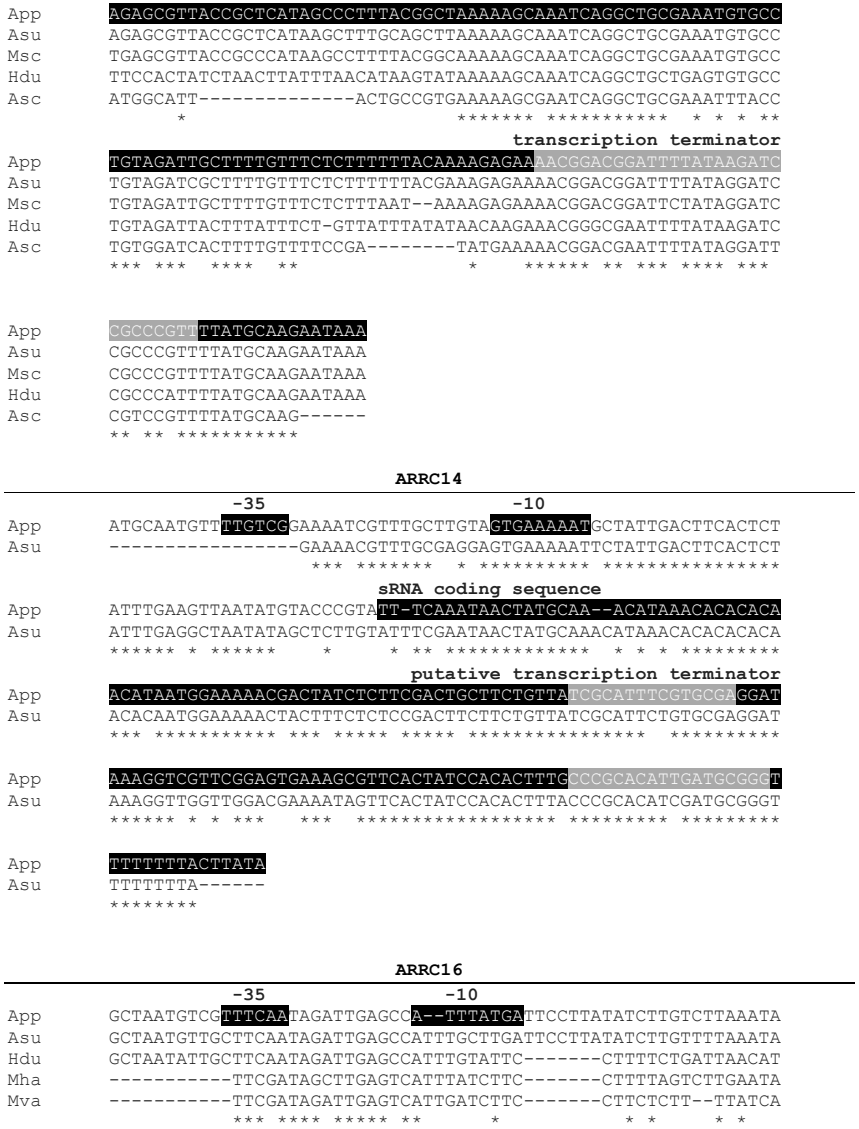

sRNA coding sequence
App AAAAGATGTTGCGGGTAGTTCACACTAATTTTATTTAACTTACAGGATACCGAGCTTGTA
Asu AAAAGATGTTGCGGGTAGTTCACACTAATTTTATTTAACTTACATGATACCGAGCTTGTA
App TGCCTAAGTCTTAGATAAGGTGTGCAAGCCTGTGGCAAGCGGTTAAAATTCACAAAATTT
Asu TGCCTAAGTCTTAGATAAGGTGTGCAAGCCTGTGGCAAGCGGTCAAAATTCACAAAATTT
putative transcription terminator
App TTGCAAAATCAGACCGCTT-ACACACCGGATAGTAAAGAAATGCACTATCCCCTCGTGTT
Asu TTGCAAAATCAGACCGCTTATCACACCGGATAGTAAAGAAATGCGCTATCCCCTCGTGTT
App TAGAAAGGTGTCAAATTCAACTTAACTTTCAGAGGGTTATGCA
Asu TAGAAAGGTGTCAAATTTAACTTAACTTTCAGAGGGTTATGCA

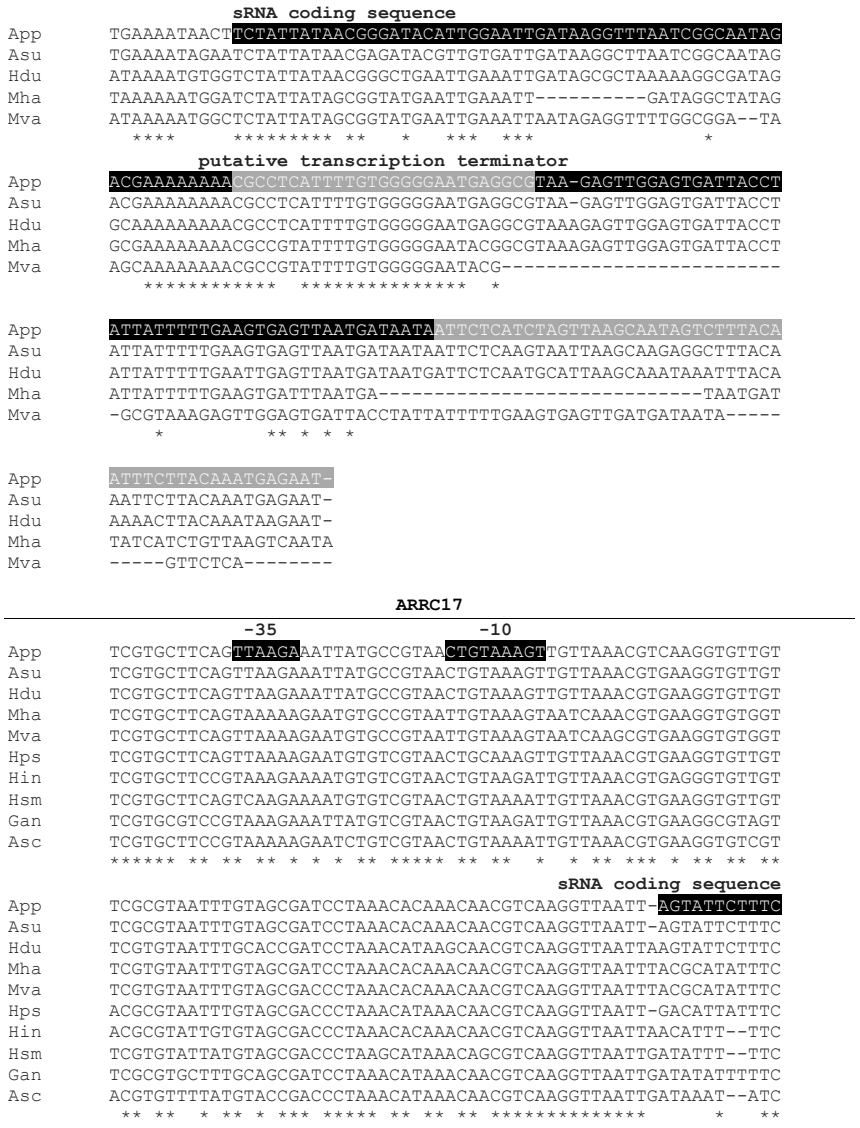

putative transcription terminator
App ---------ACCTGAAAGATTTTACCTCGTCGGTAGAGCTTATCGCTCTTCTCCAGCTACC
Asu ---------ACCTGAAAGATTTTACCTCGTCGGTAGAGTTTTCACTCTTCTCCAGCTACC
Hdu GCACCTGAAAGATTTTACCTCGTCGGTAGAGCCTAAACTCTTCT-------CCAGCTACC
Hps TGTCGCACCTGAAAGATT---TTACCTCGTCGGTGTAGGGTTAT-------TCCCCACTC
Mva -------GCACCTGAAAGAGTTTACCTCGTCGGTGAAATGTTAATTCATTTCTCTCCAGC
Mha -------GCACCTGAAAGAGTTTACCTCGTCGGTGAAATGTTAATTCATTTCTCTCCAGC
App CTTCATCCGAACGTATCTTTT
Asu TTCATCCGAAC----------
Hdu TTCATCCGAACGTTTTCTTTT
Hps TCCAGCTATCTTCATCCGAAC
Mva TACCTTCATCCGAAC------
Mha TACCTTCATCCGAAC------

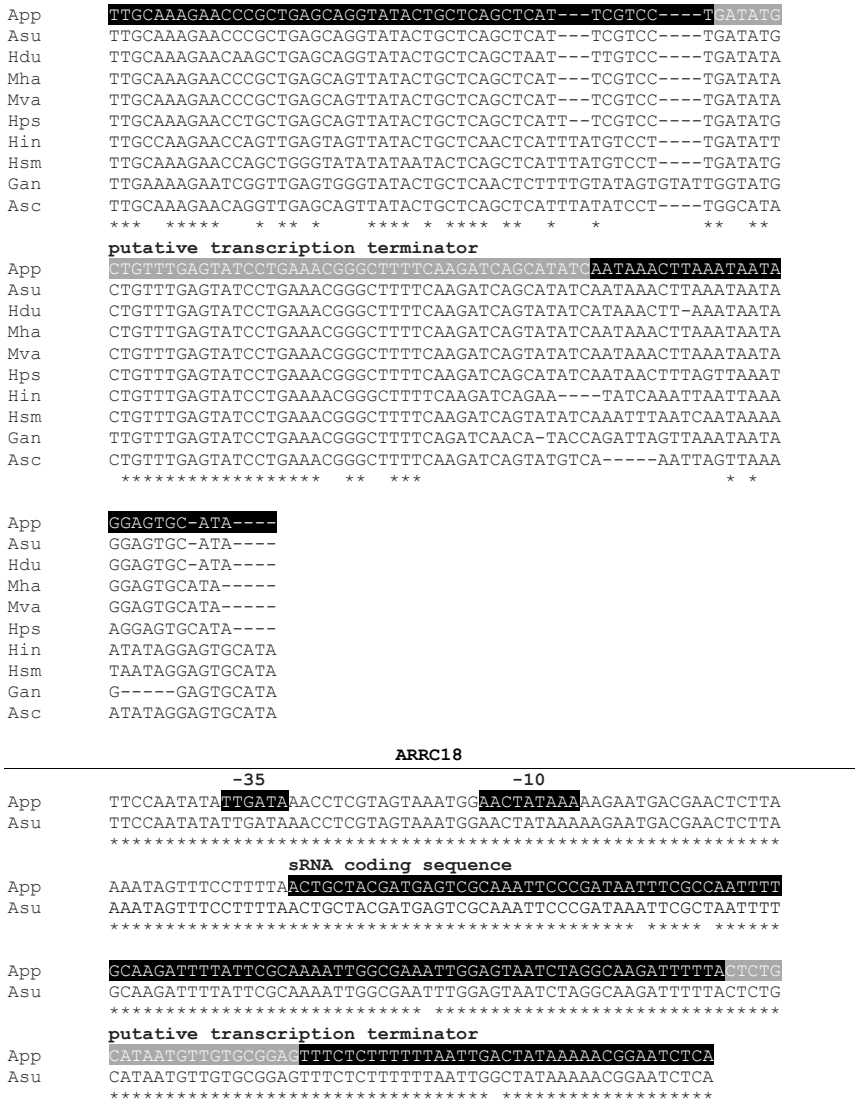

ARRC07
-35
-10
App TAAACATAACTTACAAGGAAAGAAAAGCAGGCTATAAGCGGTCGAATTTCGCCAAAAATT
Asu ----------
Hdu ----------
Hps ----------
Mva ----------
Mha ----------
sRNA coding sequence
App TGCAAAATCCTAAACACAAATAGAAAATAGGTTGTACTTTTCTCCCCTGTCCTTTTGCCT
Asu ----------ACTTTTCTCCCCTGTCCTTTTGCCT
Hdu ----------ACTTTTCTCCCCTGTCCTTTTGCCT
Hps ----------TACTTTTCTCCCCTGTCCTTTTGCCT
Mva ----------ACTTTTCTCCCCTGTCCTTTTGCCT
Mha ----------ACTTTTCTCCCCTGTCCTTTTGCCT
App GAGCGTTTCATGCTTACGCACT-------TGCGCCTTCGGCGTCCATTTTCCAT-TAACT
Asu GAGCGTTTCATGCATTCGCACT-------TGCGCCTTCGGCGTCCATTTTCCAT-T---T
Hdu GAGCGTTTCACATAAATTTGCTTTATATTTGCGCCTTCGGCGTCCATTTTCTGTACTATA
Hps GAGCGTTTCGCATTCATTGCTTT--------CGCCTTCGGCGTCCATTTT-----CTGCT
Mva GAGCGTTTCACACATTGCTGCATT-----TTCGCCTTCGGCGGCT-------------GA
Mha GAGCGTTTCACACATTGCTGTATT-----TTCGCCTTCGGCGGCG-------------GA
App ACATGGATTGGATCTCTCCAAGGGTTCGTCCAGTAACTGTCCTCGCTATTGCTAGC----
Asu ACATGGATTGGATCTCTCCAAGGGTTCGTCCAGTAACTGTCCTCGCTATCGCTAGC----
Hdu GTACAGAGTGGATCTCTCCAAGGGTTCGTCCAGTAACTGTCCTCACATAGTTAAACTATT
Hps ATGCAGAGTGGATCTCTCCAAGGGTTC--------------GTCCAATAGCTGTCCTCGC
Mva TTTTTCATCAGCTCTCTCCAAGGGTTCGTCCAGTAACAGTC-CTCGCAACTTAGTT----
Mha TTTTA-ATCCGCTCTCTCCAAGGGTTCGTCCAGTAACAGTCCCCGCAGTGTTCACT----

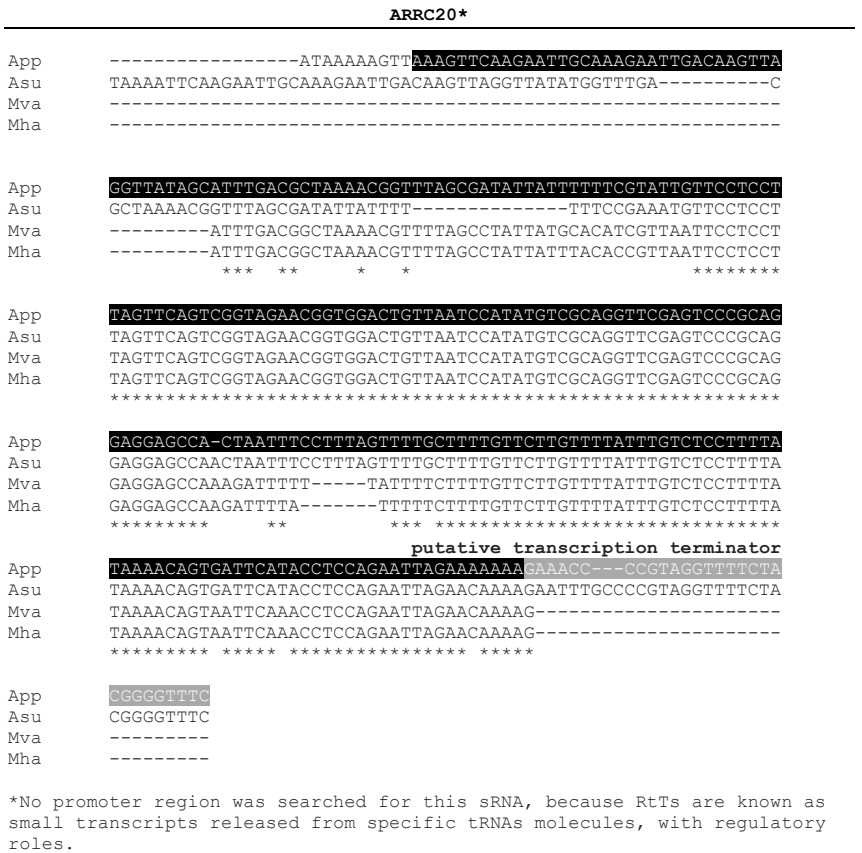

ARRC08
-35
-10
App ACCTGAACAATTAAAAGATAACTTAGTTGCGTTATTAGCGGCGTTAAACAAAGCTAAACC
Mva --CTGAACAATTAAAAGATAACTTAAACGCATTATTAGCGGCATTAACTAAAGCTAAGCC
Hdu ------ACAATTAAAAGAAAACCTTCAAGCATTGTTAGCAGCATTAAATAAAGCTAAACC
Asu ----------
Mha ----------
App GACAACAGCTAAAGGTATCTTCATCAAGAAAGTAAGCATCTCTACAACGATGGGTGCTGG
Mva GACTACAGCGAAAGGTATCTTCATCAAGAAAGTAAGCATCTCTACAACTATGGGTGCTGG
Hdu AACCACTGCGAAAGGTATCTTCATTAAGAAAGTAAGTGTCTCTACAACAATGGGCGCGGG
Asu ----------
Mha ----------
sRNA coding sequence
App TGTTGCTGTTGATCAAGCATCACTTTAATTTCTAATCGAAGTTAAAAC-TTTACAAGGT-
Mva TGTTGCTGTTGATCAAGCTTCACTTTAATTTCTAGTTAGAAGTTAAAACTTTACAGGGTT
Hdu TGTAGCTGTTGAGCAAACCTCACTTTAATACTGACTTTACAGGGT----------
Asu ----------
Mha ----------CTTTACAGGGTC
App CGTGGATTATTGTATAATTTGCGACCTTAACTTGCGAGAGCAAGCTGATGGTGCTTAGCC
Mva GCAGATTATCTGTATAATTTGCGACCTTAACTTGCGAGAGCAAGCTGATGGTGCTTAGCC
Hdu -CGTATTATCGCGTATAATTTCGACCTTAACTTGCGCGAGCAAGCTGATGGTGCTTAGCC
Asu ----------TGATGGTGC-TTGC-
Mha GCAGATTATCTGTATAATTTGCGACCTTAACTTGCGAGAGCAAGCTGATGGTGCTTAGCC
App TATCTAAGCCCCGTCCAAGACCGTAGGTGAAT-AAGTTTTCTTATTCTTAATAAAAACCT
Mva TATCTAAGCCCCGTCTAAGACTGTAGGGGTGAAAACT----------TAATAATCCT
Hdu TATCTAAGCCCCATCCAAGACCGTAGGTGGATAGCTT--GCTATTCTTAATAACAAGCCT
Asu CTATCAAGCCCCGTCCAAGACCGTAGGTGAAATGATGTTAATCATCTCTTAATTTGGCCT
Mha TATCTAAGCCCCGTCTAAGACTGTAGGGGTGAAAACT----------TAATAATCCT


putative transcription terminator
App ACGTAGATGGTGAACAGACAGAATTTTCTGCTTCTGGACACCTTAGGCTCAGAAGATTGC
Mva ACATAGATGGTGAACAGACAGAATTATCTGCTTCTGTACACCTTAGGCTCAGAAGATTG-
Hdu ACGTAGATGGTGAACAGACAGAATTATCTGCTTCTGGACACCTTAGCCTCAAGAAGATGC
Asu ACGTAGATGGTGAACAGACAGAATTTTCTGCTTCTGGACACCTTAGGCTCAGAAGATTGC
Mha ACATAGATGGTGAACAGACAGAATTCTCTGCTTCTGGACACCTTAGGCTCAAAGAAGATT
App TACTTGCGAT--------------GCAAGCGGTGAGTTTTTTGAGGTTTTTTGTAAATTC
Mva ------CAGCTTATTACAAGCGGTAATATTTTTGGGATTTTTTGTAAATTCTGCTTGAAA
Hdu CG------------TATAAGCGGTAATATTTTTGGGATTTTTTGTAAATTCTGCTCACGA
Asu TACTTGCGAT--------------GCAAGCGGTGAGTTTTTTGAGGTTTTTTGTAAATTC
Mha GCAGTTTGTAGCAATGCAAGCGGTAATATTTTTGGGATTTTTTGTAAATTCTGCTTGAAA
App CACTTCGGTGGAGTGTATCAGGAGCTAAAACCA
Mva AGGCAGAGTGTATCAGGA-----GCTAAAACCA
Hdu AAGTGGCAGAGTGTATCAGGAGCTAAAACCA--
Asu CACTTTGGTGGAGTGTATCAGGAGCTAAAACCA
Mha AGGCAGAGTGTATCAGGA-----GCTAAAACCA
ARRC09
-35
-10
App TGGTGTTTAGATGCTAAACGTTTATAAGCACGTTTAATATCGTTCTCACTTGCCCCTTTC
Hin ---TGTTTAGACGCTAAACGTTTATACGCACGTTTAATATCGTTCTCACTCGCCCCTTTT
Asu TGGTGTTTAGATGCTAAACGTTTATAAGCACGTTTAATATCGTTCTCACTCGCCCCTTTC
sRNA coding strand
App TGTAAGCCAAGGACTTCGTAGTAATCTTTTTTTGCCATAGTGTTTCGTTTGTAAAATTTT
Hin TTCAATCCAAGGACTTCGTAGTAATCTTTTTTTGCCATAGTGTTTCGTTTGTAAATTTTA
Asu TGTAAGCCAAGGACTTCGTAGTAATCTTTTTTTGCCATAGTGTTTCGTTTGTAAAATTTT
putative transcription terminator
App ACGGAAATTTAACCGCTTGTCATGAACTCCCCTCTTTAGCAA----------
Hin TAGA-AATTTGACCGCTTGTAACATACTCC----------CCTCT
Asu ACGGAAATTTGACCGCCTGTCATAAACTCCCCTCTTTAGCAAAGAGGGGAACTCCCCTCT
App -------AGAGGGGCAGGGGGAGATTTGGCAGAAGTAACTTTAGCTCATAAGAGAATTTG
Hin TTAACAAAGAGGGGCGGGGGGAGATTTGGCAGAAGTAAATTTAGCTCATAAGAAAATTTG
Asu TTAGCAAAGAGGGGCAGGGGGAGATTTGGCAGAATAGCATTACGCTCATACGAAAATTTG
App ACATCGTTACCAAATCTCCCCTAACCCCTCTTTTCTAAAGAGGGGGACTTGTT
Hin ATATCGCTATCAAATCTCCCCTAACCCCTCTTTTCTAAAGAGGGGGA------
Asu ACCTTGTCATCAAATCTCCCCTAACCCCTCTTTTCTAAAGAGGGGGAC-----
ARRC11
-35
-10
App GTTTCAGCAATTGAATATCCTTAAATGATAGGTTTTAGGGGGATAAAAAAATTATTTTTA
Asu -----AAAAT----------TTTTAAA
sRNA coding sequence
App TTTTTTTAGAAATTGTCCAATAAATAGGCTTCCCAATTATGTAAAAAGTGCGTATCATAC
Asu TATTTTTAAAAATTGTCGAATAAATAGACTTCCAAAATAGGAAAAAAGTGCGTATCATAC
transcription terminator
App GCACCTCA-AAATCAGCTGAAGCTGACATAGATTGATTAAAGAGTACGGTTTTTTAGCCG
Asu GCACCTCGAAAATCAGGTTAGCCTGAAAGGTTTAAATAAGAAAGTACGGTTTTTTAGCCG
App TACTTTTTATTTGGATAGTTATCTCGTTAAATTTGTTTTAAATTTAATCACTTACGGAAG
Asu TACTTTTTATTTGGATGTCTCTCTCGTTAAAATCGCTTTAATTTTAATCATTTACGG---
App A
Asu -
ARRC12
-35
-10
App GTAAGTGACGTTTACGTTTAGTGGTTTTCTTAGTCAAAATATGACGTAAGTGAGATTGTT
Asu GTAAGTGACGTTTACGTTTAGTGGTTTTCTTAGTCAAAATATGACGTAAGTGAGATTGTT
Msc GTAAGTGACGTTTACGTTTAGTTGTTTTCTTAGTCAAAATATGACGTAAGTGAGATTGTT
Hdu -TAAATGACGTTTACGTTTAGTGGTTTTCTTAGTCAAGATATGACGTAAGTGAGATTGTT
Asc -----TGACGTTTACGTTTAGTGGTTTTCTTAGTCAAAATATGACGTAAGTGAGATTGTT
App TACGTTTGAAACCGCCTGAAGCTGTTTTCTTGAAACGCTTAGCAGCACCACGTACTGTTT
Asu TACGTTTGAAACCGCCTGAAGCTGTTTTCTTGAAACGCTTAGCAGCACCACGTACTGTTT
Msc TACGTTTGAAACCGCCGGAAGCCGTTTTTTTGAAACGCTTCGCAGCACCACGTACTGTTT
Hdu TACGCTTGAAACCGCCTGAAGCTGTTTTCTTAAAACGCTTAGCAGCACCACGTACTGTTT
Asc TACGTTTGAAACCGCCGGAAGCTGTTTTTTTGAAACGCTTAGCAGCACCACGTACTGTTT
sRNA coding sequence
App TAATTTTAGGCATTGTTTTATAAACTCCGCATTGTTTTATCTAATACATAATAATCAGGC
Asu TAATTTTAGGCATTGTTTTATAAACTCCGCATTGTTTTACATAATACACAATAATTAGGC
Msc TAATTTTAGGCATTGTTAAAAAAACTCCGCATTGTTTTGTTTGATACACAATAATCAGGC
Hdu TAATTTTAGGCATTGTTTTAGAAACTCCGCATTGTTTGTTAATCAAAAGTCAATTA-GGC
Asc TAATTTTAGGCATTGTTTAAAAACTCCGCATTGTTT--TGTTAACACTTGATAGTCAGGC
App GAAAAATACGCTTATCGCTAACCGTCTTTTTACTTGC--AAAGCACTAATTATCTCGAAC
Asu GAAAAATGCGATTATTGCTAATCATATTTTTACTTGG--AAAGCGCTAATTATCTCGAA-
Msc GAAAAATACCGCCATTTCTGACCGCACTTTTACTTGG--AAAGCACTAATTATCTCTGAT
Hdu GAACAGTATTACTAATGTTACTTGGAAGCAACT------AAT-GACATATAAAAAATAGT
Asc GAAAAACACGATTATTTTCAACCGCACTTTTGCTTTTGTAAAGCACTGCTTATCTCTCAA


transcription terminator
ARRC14
-35
-10
sRNA coding sequence
putative transcription terminator
ARRC16
-35
-10


sRNA coding sequence
putative transcription terminator
ARRC17
-35
-10
sRNA coding sequence





*No promoter region was searched for this sRNA, because RtTs are known as small transcripts released from specific tRNAs molecules, with regulatory roles.

ARRC21*

```
App      AAAAAGCACTTGACCATTTTGTTTAAATCCGTATTATATGCGCCTGTTACGCAACGTTAA
Asu      AAAAAGCACTTGACCATTTTTTTTAAATCCGTATTATATGCGCCTGTTACGCAACGTTAA
Hdu      AAAAAGCACTTGACGCTCTTACTAAAATCCGTATTATACACGCCTGTTATGCAATATTAA
Hps      ------------------------AAATCCGTATGATACGCCCTCGTTACGCAACGTTAA
Mva      ------------------------AAATCCGTATGATATGCCCTCGTTACGCAAGATTAT
Mha      ------------------------AAATCCGTAAGATATGCCCTCGTTACGCAAGATTAT
Agp      ----------------------AAATCAGTATTATAAGCGCTCGTTGTTAAATGTTACTT
Gan      ------------------------------------------------------------
Hsm      -------------------------GTAATA------TATCCGCTCGTTACAAACAATGA
Agc      ----------------------AAATCAGTATTATAAGCCCTCGTTGTTAGATGTT-AAC
Asc      ------------------------------------------------------------
Msc      ------AAATCC-------------GTATTATAAGCACCCGTTACACAGCGTAACCTTGT
Pmu      ------------AAATCAGTAT------TATAAGCCTCCGTTACGCAATGATATGCAAAT

App      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Asu      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Hdu      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Hps      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Mva      GGGTCATTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Mha      GGGTCATTAGCTCAGCCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Agp      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Gan      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Hsm      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Agc      GGGTCGTTAGCTCAGCCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Asc      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
Msc      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGT------
Pmu      GGGTCGTTAGCTCAGTCGGTAGAGCAGCGGACTTTTAATCCGTTGGTCGAAGGTTCGAAT
         ***** ********* ************************************
```

putative transcription terminator

```
App      CCTTCACGACCCACCATTTAACTTGCAACGCCTTCTAAAGGGTCGTTA-GC------TCA
Asu      CCTTCACGACCCACCATTTAACTTGCAACACCTTCTAAAGGGTCGTTA-GC------TCA
Hdu      CCTTCACGACCCACCACTTAATAATTTGCACACCTACATTCTAAAGGG-TCGTTAGCTCA
Hps      CCTTCACGACCCACCACTTAATCAGCGTTCTTTTCTAAAGGGTCGTTA-GC------TCA
Mva      CCTTCATGACCCACCATTTAATCTTCAAAACACCTCTTTGGGTCATTA-GC------TCA
Mha      CCTTCATGACCCA-CCATTAATCTTCAAAACACCTCTAAGGGTCATTA-GC------TCA
Agp      CCTTCACGACCCACCAATTTAACAATATACCCCTTCAAAGGGTCGTTAG-------CTCA
Gan      CCTTCACGACCCACCATTTATTATAATACCCTTATGGGTCGTTAGCTC-AGTCGGTAGAG
Hsm      CCTTCACGACCC----ACCATTTTATTTGTAATATCCCAATTTATGGG-TCGTTAGCTCA
Agc      CCTTCACGACCCACCACTTTTAAACAATA----TATCCCAAATAATGGGTCGTTAGCTCA
Asc      CCTTCACGACCCACCACTTAAATCTGGTTTACCCATATCAGTATGGGT-CGTTAGC-TCA
Msc      --------------------TCGAATCCTTCACGACCCACCACTTAAA-ATTTGGT-TTT
Pmu      CCTTCACGACCCACCACTTAA-----------------------------------------

App      GTCGGTAGAGCAG----------------------------
Asu      GTCGGTAGAGCAG----------------------------
Hdu      GTCGGTAGAGCAG----------------------------
Hps      GTCGGTAGAGCAG----------------------------
Mva      GTCGGTAGAGCAG----------------------------
Mha      GTCGGTAGAGCAG----------------------------
Agp      GTCGGTAGAGCAG----------------------------
Gan      C----------AG----------------------------
Hsm      GTCGGTAGAGCAG----------------------------
Agc      GTCGGTAGAGCAG----------------------------
Asc      GTCGGTAGAGCAG----------------------------
Msc      ATCCATATTCAATATGGGTCGTTAGCTCAGTCGGTAGAGCAG
Pmu      -----------------------------------------
```

*No promoter region was searched for this sRNA, because RtTs are known as small transcripts released from specific tRNAs molecules, with regulatory roles.