The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages

Mobarak A. Mraheil[1], André Billion[1], Walid Mohamed[1], Krishnendu Mukherjee[2], Carsten Kuenne[1], Jordan Pischimarov[1], Christian Krawitz[1], Julia Retey[3], Thomas Hartsch[3], Trinad Chakraborty[1,*] and Torsten Hain[1,*]

[1]Institute of Medical Microbiology, Justus-Liebig-University, Frankfurter Strasse 107, [2]Institute of Phytopathology and Applied Zoology, Justus-Liebig-University, Heinrich-Buff-Ring 26-32, 35392 Giessen and [3]Genedata Bioinformatik GmbH, Lena-Christ-Strasse 50, D-82152 Planegg-Martinsried, Germany

ABSTRACT

Small non-coding RNAs (sRNAs) are widespread effectors of post-transcriptional gene regulation in bacteria. Currently extensive information exists on the sRNAs of *Listeria monocytogenes* expressed during growth in extracellular environments. We used deep sequencing of cDNAs obtained from fractioned RNA (<500 nt) isolated from extracellularly growing bacteria and from *L. monocytogenes* infected macrophages to catalog the sRNA repertoire during intracellular bacterial growth. Here, we report on the discovery of 150 putative regulatory RNAs of which 71 have not been previously described. A total of 29 regulatory RNAs, including small non-coding antisense RNAs, are specifically expressed intracellularly. We validated highly expressed sRNAs by northern blotting and demonstrated by the construction and characterization of isogenic mutants of *rli31*, *rli33-1* and *rli50** for intracellular expressed sRNA candidates, that their expression is required for efficient growth of bacteria in macrophages. All three mutants were attenuated when assessed for growth in mouse and insect models of infection. Comparative genomic analysis revealed the presence of lineage specific sRNA candidates and the absence of sRNA loci in genomes of naturally occurring infection-attenuated bacteria, with additional loss in non-pathogenic listerial genomes. Our analyses reveal extensive sRNA expression as an important feature of bacterial regulation during intracellular growth.

INTRODUCTION

The availability of an increasing number of complete bacterial genome sequences along with recent technical advances in DNA sequencing has led to an explosion in the identification of numerous small, non-coding RNAs (sRNAs) (1–3) and this number is constantly growing (2,4–6). A combination of both computational and novel experimental approaches have demonstrated the ubiquity of sRNAs and has led to the description of many functionally important sRNAs in organisms ranging from eubacteria to humans (7–12).

The majority of sRNA candidates identified to date are thought to regulate gene expression by hybridizing with target mRNA thus modulating its stability and/or translation activity. In addition, some sRNAs bind to proteins and modulate their activity or build functional complexes (13,14). In bacteria, the main function of sRNAs are in coordinating adaptation to environmental changes and signals by controlling target gene expression and includes, for example, responses to iron limitation, oxidative stress and low temperature (4,13–15).

The role of sRNA in controlling virulence and pathogenesis has been demonstrated for a number of Gram-negative bacteria including *Escherichia coli*, *Pseudomonas aeruginosa*, *Salmonella typhimurium*, *Vibrio cholerae*, *Chlamydia trachomatis* and *Helicobacter pylori*, as well Gram-positive bacteria such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Clostridium perfringens* (16–18).

*To whom correspondence should be addressed. Tel: +49 641 99 46400; Fax: +49 641 99 46409; Email: torsten.hain@mikrobio.med.uni-giessen.de
Correspondence may also be addressed to Trinad Chakraborty. Tel: +49 641 99 41251; Fax: +49 641 99 41259;
Email: trinad.chakraborty@mikrobio.med.uni-giessen.de

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

An initial approach to sRNAs detection was the characterization of those RNAs which bind to Hfq, a protein originally identified as a host factor needed for Qβ bacteriophage replication in *E. coli* (19,20). In *E. coli*, Hfq modulates the activity of several sRNAs and acts as an RNA chaperone to promote sRNA–mRNA duplex formation (21,22). Analysis of *hfq* mutants of *P. aeruginosa*, *V. cholerae*, *Legionella pneumophila* and *Brucella abortus* revealed attenuated growth in macrophages or mice (23–26).

In the Gram-positive, facultative intracellular pathogen bacterium *L. monocytogenes*, it was shown that Hfq contributes to pathogenesis in mice (21,27). Presently, three sRNAs have been identified by co-immunoprecipitation using Hfq (28). Additional studies have led to the description of further 12 sRNAs including the σ^B-dependent SrbA transcript (29,30). Recently a genome-based tiling array study provided a catalog of sRNA candidates expressed under different physiological growth conditions (31). These data have largely been confirmed by a study in which *L. monocytogenes* and its isogenic SigB mutant strain were grown under stationary growth conditions and the repertoires of RNA produced examined by deep RNA sequencing (32).

Listeria monocytogenes is the causative agent of listeriosis, a severe human infection with a high mortality rate. The bacterium inhabits numerous ecological niches, it can multiply at high salt concentrations (10% NaCl) and broad ranges of pH (4.5–9) and temperature (0–45°C) (33). A hallmark of this human pathogen is its ability to invade and survive inside vertebrate and invertebrate host cells, wherein the bacterium can freely multiply within the cytosol and can induce actin-based movement. Actin-based movement allows the bacterium to spread from cell-to-cell which leads to fatal outcomes of listerial infection. Prior to infection, internalin A and B induce the first step of the infection process in non-phagocytic cells by interacting with the eukaryotic host cell and promote the intracellular uptake of the pathogen after binding with the E-Cadherin and c-Met receptors in mammals. However, the main virulence genes, responsible for the intracellular life cycle of *L. monocytogenes* are clustered in a ∼9 kb chromosomal region. This virulence cluster encodes the genes *viz*., listeriolysin O (LLO) and phosphatidylinositol phospholipase A (PlcA), which are responsible for the escape of bacterium from the primary vacuole. Another phospholipase (PlcB) is required for bacterial cells to escape from the secondary vacuole and is processed with the assistance of the metalloprotease (Mpl). To facilitate actin recruitment for intracellular movement, the bacterium modulates components of the host cell machinery with the surface protein ActA, which is anchored within the cytoplasmic membrane. Finally, the regulation of the virulence gene cluster is dependent on the master regulator PrfA, a Crp/Fnr-like transcriptional regulator (34,35). In this context it is reasonable to assume that sRNAs may play a versatile role in the adaptation mechanisms of *L. monocytogenes* to these different environments.

Currently, there is extensive information available regarding the transcriptome of *L. monocytogenes* when grown under conditions that are external to the host cell (31,32). These include data for bacteria grown extracellularly in broth (exponential and stationary growth) under different conditions of stress, including low oxygen, low temperature (30°C), in blood and the lumen of the infected gut as well as analysis of several isogenic mutants such as Δ*prfA*, Δ*sigB* and Δ*hfq*. However, information as to whether specific sRNAs are expressed during intracellular growth and their respective roles are largely unknown (11).

Here we describe a genome-wide search for sRNAs expressed in the wild-type of *L. monocytogenes* during extracellular and intracellular growth. Extracellular cultures were grown until exponential phase like previous studies (31,32). Intracellular cultures were grown in P338D1 murine macrophages. RNA was collected 4 h post-infection and size-fractioned to <500 nt. cDNA generated from size-fractioned RNA was deep sequenced and provided a comprehensive view of listerial sRNA candidates preferentially induced following infection in murine macrophages. We report on the discovery of 150 putative regulatory RNAs of which 29 are specifically expressed intracellularly. Analysis of several sRNA candidates highly expressed during intracellular growth revealed that these loci are highly conserved in pathogenic *L. monocytogenes* strains. Isogenic mutants lacking these loci exhibit attenuated virulence in *in vivo* models of infection.

MATERIALS AND METHODS

Bacterial strains and growth conditions

The strain *L. monocytogenes* EGD-e (36) and its chromosomal deletion mutants (Supplementary Table S1A) were used in this study. Bacteria were grown in BHI broth (VWR) overnight at 37°C with shaking at 180 r.p.m. (Unitron, Infors) until mid-exponential phase (OD$_{600\ nm}$ 1.0). Overnight cultures were diluted 1:50 in 20 ml fresh BHI broth using a 100 ml Erlenmeyer flask and were incubated at the same conditions mentioned above until OD$_{600\ nm}$ 1.0.

RNA isolation

For RNA isolation from *L. monocytogenes* grown extracellularly in BHI broth (VWR) until mid-exponential phase, aliquots of 0.5 ml bacterial culture were treated with 1.0 ml RNA protect (Qiagen) for 5 min, the bacterial cells were collected by centrifugation for 10 min (8000*g*) and subsequently stored at −80°C until use. RNA extraction from intracellularly grown *L. monocytogenes* in macrophages 4 h post-infection was performed as described previously (37). Briefly, infected host cells were lysed using cold mix of 0.1% (w/v) sodium dodecyl sulfate, 1.0% (v/v) acidic phenol and 19% (v/v) ethanol in water. The bacterial pellets were collected by centrifugation for 3 min (16 000*g*).

Total RNA was extracted using miRNeasy kit (Qiagen) with some modifications. The collected pellets were washed with SET buffer [50 mM NaCl, 5 mM EDTA and 30 mM Tris–HCl (pH 7.0)]. After centrifugation at 16000 *g* for 3 min pellets were resuspended into 0.1 ml

Tris–HCl (pH 6.5) containing 50 mg/ml lysozyme (Sigma), 25 U of mutanolysin (Sigma), 40 U of SUPERase (Ambion), 0.2 mg of proteinase K (Ambion) and incubated at 37°C for 30 min at 350 r.p.m. QIAzol (Qiagen) was added, mixed gently and incubated for 3 min at room temperature. An additional incubation at room temperature was done after adding 0.2 volume chloroform followed by centrifugation at 16 000*g* at 4°C for 15 min. The upper aqueous phase, containing RNA, was transferred to a new collection tube and 1.5 v of 100% ethanol was added and mixed thoroughly. The probes containing RNA and ethanol were transferred into columns supplied with the miRNeasy Kit (Qiagen) and treated according to the manual including an on-column DNase digestion (RNase-Free DNase, Qiagen). RNA was eluted by RNase-free water and stored at −80°C until needed. The quantity of the isolated total RNA was determined by absorbance at 260 and 280 nm, and the quality was assessed using Nano-chips for Agilent's 2100 Bioanalyzer. For detection and estimation of the small RNA fraction within the isolated total RNA, a small RNA-chip (Agilent) was used, which visualizes RNAs with sizes ranging from 20 to 150 nt.

RNA sequencing and data analysis

The fraction of the total RNA <500 bp was used for RNA-Seq as described below. The solubilized RNAs were first treated with tobacco acid pyrophosphatase (Epicentre) as recommended by the manufacturer. This treatment allows discriminating primary 5′-ends generated by transcription initiation from 5′-ends generated by RNA processing. Then the small RNAs were poly(A)-tailed using poly(A) polymerase followed by ligation of a RNA adapter to the 5′-phosphate of the RNA. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and M-MLV reverse transcriptase (Promega) assay. The resulting cDNAs were PCR-amplified in 20 cycles to a concentration of 20 ng/µl using a high fidelity DNA polymerase. For size fractionation PCR-amplified cDNA was run on a preparative 6% PAA-gel. PCR products containing small RNA sequences of 20–500 nt were isolated from PAA-gel. The cDNAs were purified using the Macherey and Nagel NucleoSpin Extract II kit. 454 pyrosequencing using GS FLX Titanium series chemistry (Roche) were carried out by Eurofins MWG Operon (Germany).

A sequence file was created for each condition, respectively. In total 189 381 reads with approximately 31 million bases were analyzed. After clipping of 5′-linker and poly(A)-tail all reads shorter than 16 nt were removed. Resulting reads were further processed via a hitherto unpublished software *sncRAS* (A. Billion, manuscript in preparation) which in general requires no clipping, which makes no significant difference during the mapping step, except for short reads combined with a poly(A)-tail which may lead to a deletion of this sequence. The remaining reads were mapped with NCBI BLASTN 2.2.17 (38) against *L. monocytogenes* EGD-e genome (GenBank Accession: NC_003210) with an e-value of 0.001, default word size and rewards for a

nucleotide match had been set to 2. Additionally, nucleotide identity was required to be >60% combined with coverage of 80% between query and subject sequence. Sequencing reads which did not fulfil these requirements were not taken into account. Reads ranged in size from 21 to 521 nt and averaged in length in both conditions around 74 nt. 28–49% of used reads matched perfectly to the genome. An additional 4–9% of remaining reads contain one mismatch.

sRNA detection

sRNA detection was carried out by the software *sncRAS* as well. The first step removes transcripts of rRNA and tRNA, plus all reads from intergenic regions (IGR) of rRNA and tRNA genes. The second step identifies locations for potential sRNAs. Regulatory RNAs were grouped in three general classes: sRNAs, antisense small non-coding RNAs (asRNAs) and riboswitches including *cis*-regulatory RNAs (Supplementary Table S3). Valid reads, which mapped within an IGR and fulfilled the cut-off criteria, were used to identify a pile-up of sequences which combined have a minimal length of 50 nt and do not overlap with adjacent genes. Furthermore putative sRNA candidates were required to have ≥10 reads in one condition. For asRNA detection similar criteria were applied, however a 5′- or 3′-overlap of its possible target gene was allowed. The resulting candidates from automatic classification were re-evaluated by visual inspection with the Integrative Genomics Viewer (IGV) version 1.4.2 (http://www.broadinstitute.org/igv). Detection of *cis*-regulatory RNAs was performed with manually evaluated candidates by searching against the Rfam 9.1 database (39). All hits with a covariance model score of more than 50 were classified as *cis*-regulatory RNAs including riboswitches.

Transcription factor binding site analysis

In order to identify putative regulatory motifs within regulatory RNAs candidate promoter regions, a list of known consensus binding boxes was compiled from the literature (Supplementary Table S4). For each putative regulatory sRNA, a region of 50 nt upstream was scanned for transcription factor binding sites. Most binding sites were detected using regular expression search (FUZZNUC, EMBOSS package) allowing one mismatch. Sigma70 promoters were identified using HMMER3 (40) on six different HMMs from the prokaryotic promoter prediction site (http://bioinformatics.biol.rug.nl/websoftware/ppp/).

Terminator identification

Terminators were identified using a pre-calculated prediction of TransTermHP2.0 (41). All terminator-like hairpins with confidence score ≥30 were taken into account. Sliding windows of ±10 nt around the 3′-end of each putative regulatory RNA were used to annotate putative terminators which additionally required at least one base overlap with this defined area. If several candidates within the search window were found, the one located closest to the 3′-end was used.

Small open reading frame detection

Small open reading frame (ORF) detection for *L. monocytogenes* (36) was based on predictions from GenDB (42) as described previously for *L. welshimeri* (43) and *L. seeligeri* (44) applying a lowered ORF detection cut off (>10 amino acids).

Data visualization

For visualization BLAST results were pre-processed by SAMtools package version 0.1.7 (45) and uploaded into the IGV version 1.4.2 (http://www.broadinstitute.org/igv) for further manual inspection. Circular chromosomes were precalculated and visualized by GenomeViz 1.3 (46). Venn diagrams were drawn by VENNY (http://bioinfogp.cnb.csic.es/tools/venny/index.html).

Northern blot analysis

For northern blotting, 15 μg of total RNA was separated on 10% polyacrylamide gels containing 7 M urea. RNA was transfered to Biodyne nylon transfer membrane (VWR) via electroblotting and UV cross linked. Oligodeoxynucleotides (Supplementary Table S1B) were end-labeled with γ^{32}-P-ATP (Hartmann Analytik) using T4 polynucleotide kinase as recommended by the manufactured protocol (Fermentas). Pre-hybridization was performed for 3 h in DEPC-treated water containing 0.1 mg/ml salmon sperm DNA, 6 × SSC, 0.5% SDS and 2.5 × Denhardt's solution. Hybridization was performed overnight after adding γ^{32}-P-ATP end-labeled oligodeoxynucleotide probes to northern membranes under the same conditions used for pre-hybridization. The hybridization temperature for oligodeoxynucleotides ranged between 50 and 65°C. Membranes were exposed on phosphor imaging screens (GE Health Care) and analyzed with Typhoon 9200 (GE Health Care). To evaluate the sizes of sRNAs we constructed a ladder containing different PCR fragments between 150 and 550 nt (Supplementary Table S1B).

sRNA mutant construction

To investigate the role of sRNA in intracellular survival *in vitro* and *in vivo*, chromosomal deletion mutants were generated. Primer sequences used to generate the isogenic mutants are presented in Supplementary Table S1B. Chromosomal deletion mutants were constructed by generating the 5′ (with primers '1-for' and '2-rev') and the 3′ (with primers '3-for' and '4-rev') flanking regions of the sRNA concerned. Construction of the fusion construct was achieved by SOEing as described earlier (47). Generation of the deletion mutants was performed as described previously (48). The deletions in the targeted genes were confirmed by PCR with according primer '5-for' and '6-rev' after verifying the sequence by automated DNA sequencing.

The complementation of the Δ*rli31*, Δ*rli33-1* and Δ*rli50** deletion mutants was carried out using the *L. monocytogenes* site specific phage integration vector pPL2 (49). Primer sequences used for the complementation of the deleted genes are listed in Supplementary Table S1B.

In vitro infection experiment

Infection experiments with *L. monocytogenes* EGD-e and its sRNA deletion mutants were performed using P388D1 murine macrophage as described previously (37).

Galleria mellonella infection model

Bacterial inoculums of wild-type *L. monocytogenes* EGD-e and its sRNA deletion mutants were injected dorsolaterally into the hemocoel of last instar larvae using 1 ml disposable syringes and 0.4 × 20 mm needles mounted on a microapplicator as described previously (50). After injection, larvae were incubated at 37°C. Larvae were considered dead when they showed no movement in response to touch. No mortality of *Galleria* larvae were recorded when injected with 0.9% NaCl. For each experiment 20 animals per sRNA deletion mutant were used and three biological replicates were performed independently.

Murine infection model

Six- to 8-week-old female BALB/c mice, purchased from Harlan Winkelmann (Borchen, Germany), were used in all experiments. *In vivo* growth kinetic and survival of wild-type *L. monocytogenes* EGD-e and its sRNA mutants were tested in a mouse infection model. Infection was performed by intravenous injection of approximately 2000 viable bacteria in a volume of 0.2 ml of PBS. After 3 days, bacterial growth in spleens and livers was determined by plating 10-fold serial dilutions of organ homogenates on BHI. The detection limit of this procedure was 10^2 CFU per organ. Colonies were counted after 24 h of incubation at 37°C.

Ethics statement

This study was carried out in strict accordance with the regulation of the National Protection Animal Act (§7-9a Tierschutzgesetz). The protocol was approved by the local Committee on the Ethics of Animal Experiments (Regierungsbezirk Mittelhessen) and permission was given by the local authority (Regierungspraesidium Giessen, Permit Number: GI 15/5-Nr.63/2007).

Statistical data analysis of infection experiments

All infection experiments were performed a minimum of three times. Significant differences between two values were compared with a paired Student's *t*-test. Values were considered significantly different when $P < 0.05$.

Comparative genomics

In order to assess the dissemination of regulatory RNAs across the genus *Listeria* the GenBank formatted files of four pathogenic (*L. monocytogenes* 1/2a EGD-e, *L. monocytogenes* 4b F2365, *L. monocytogenes* 1/2a 08-5578, *L. monocytogenes* 1/2a 08-5923, *L. monocytogenes* 4a HCC23) and three apathogenic strains (*L. innocua* 6a Clip11262, *L. welshimeri* 6b

SLCC5334, *L. seeligeri* 1/2b SLCC3954) were retrieved from the GenBank repository (http://www.ncbi.nlm.nih .gov/genbank/index.html). These genome data as well as all candidate regulatory RNAs were then introduced to the sRNAdb database (J. Pischimarov, manuscript in preparation) which employs BLASTN to find sequence similarity. Using a cutoff of 60% nucleotide identity and 80% coverage homologs of all candidate regulatory sRNAs were identified in the aforementioned genome sequences.

RESULTS

RNA sequencing

To investigate the intracellular sRNA transcriptome profile of *L. monocytogenes* EGD-e, total RNA was isolated from bacteria grown extracellularly in BHI or from the cytosol 4 h post-infection in P388D1 murine macrophages. Following size-fractionation, cDNA with a size <500 nt was used for the cDNA sequencing studies. Samples were subjected to 454-based pyro-sequencing and base-called reads from two cDNA libraries generated from total RNA of either extracellularly (EC) or intracellularly (IC) grown *L. monocytogenes* were analysed in this study. These libraries generated a total 189 381 reads of which 116 158 (61%) were derived from RNA isolated from intracellular bacteria and 73 223 (39%) from bacteria grown in broth cultures. An overview of the filtered and mapped sequencing reads is given in Supplementary Table S2.

Sequencing reads were mapped to the genome of *L. monocytogenes* using BLASTN with an e-value of 0.001 and default word size with rewards for a nucleotide match that had been set to two. Additionally nucleotide identity was required to be >60% combined with coverage of 80% between query and subject sequence. Reads that did not fulfil these requirements were removed from the dataset. After clipping, additional linker removal, quality control and mapping against the genome of *L. monocytogenes* EGD-e, 114 459 unique reads with at least 80% sequence identity with 80% coverage and a minimum length of 21 nt remained for detailed analysis. We observed that ~49% of the IC reads perfectly match the genome with 100% sequence identity compared with only 28% of the reads from the EC cDNA library. The 'intergenome' of *L. monocytogenes,* i.e. the non-coding sequence between annotated ORFs, comprises ~10% (~300 000 bp) of the entire genome. In the intergenome fraction we observed expression of nearly one-third under intracellular condition which additionally shows the importance of the intergenome (Figure 1A). Approximately 60% of all sequence reads mapped to annotated rRNA and tRNA genes (Figure 1B and C).

Identification of sRNAs, asRNAs and *cis*-regulatory RNAs including riboswitches

Bioinformatics analysis using previously described methods ('Materials and Methods' section) for analysing RNA-Seq data for sRNAs in *L. monocytogenes* EGD-e identified 150 putative regulatory RNAs expressed under either or both growth conditions (Figure 1A and Supplementary Table S3). The putative candidates were placed into three different classes: class I comprises sRNAs, which are located within IGRs without overlapping adjacent genes; class II includes antisense RNAs, referred as asRNAs, located antisense to an annotated ORFs and class III comprises *cis*-regulatory RNAs including riboswitches.

A total of 121 regulatory RNA elements were mapped within IGRs of the *L. monocytogenes* chromosome and included 88 sRNA and 33 *cis*-regulatory RNAs including riboswitches. In addition, we detected 29 putative asRNAs (Figure 1A and Supplementary Table S3). Of the 150 candidates determined, 142 were intracellularly expressed and 121 were transcribed extracellularly (Figure 1A and Supplementary Table S3). A common set of 113 regulatory RNAs are detected in both conditions. Of the 142 candidates expressed intracellularly ~20% (29 candidates) were specific for intracellular growth and 35% (50 candidates) showed enhanced expression which was evident through an increased cDNA read number (IC/EC ratio >2) under intracellular growth conditions. In contrast ~7% (8) of the 121 extracellularly expressed transcripts were specific for extracellular growth, while 49% (60) showed enhanced expression which was evident through an increased cDNA read number (EC/IC ratio >2) under extracellular growth conditions. The sizes of the putative regulatory RNAs ranged from 50 to 517 nt with a minimum of 10 reads and a maximum read count of 1603 in at least one growth condition.

Of the 88 class I sRNAs, 85 showed intracellular, and 77 extracellular, expression. Eleven sRNAs are specific for intracellular growth whereas only three sRNAs are specific to the extracellular lifecycle. The overall numbers of reads covering the 11 intracellular specific sRNAs was in general at least 25% higher than that of the 113 common regulatory RNAs which were found in both conditions.

The class II asRNA family revealed an even higher variation in expression than the sRNAs. Of a total of 29 asRNAs detected, approximately twice that number i.e. 25 asRNAs were expressed intracellularly as compared with only 11 extracellularly. Of these 18 were specifically expressed intracellularly as compared with only 4 for the extracellular growth.

The third class of sRNAs showed no major differences in occurrence between the growth conditions. We identified 33 *cis*-regulatory RNAs in total (32 IC versus 33 EC) with only one candidate being specifically expressed extracellularly. The latter one was identified as a lysine riboswitch and is located upstream of a gene encoding an amino acid permease. However it was remarkable that the geometric mean of extracellularly expressed *cis*-regulatory RNA is at least twice as high as the one of those expressed under intracellular growth conditions.

Identification of eight major transcription factor binding sites (Supplementary Table S4) in the candidate pool suggests regulation of 24% of all candidates. A putative housekeeping promoter was found for 36 of 142 intracellular and 39 of 121 extracellular transcripts.
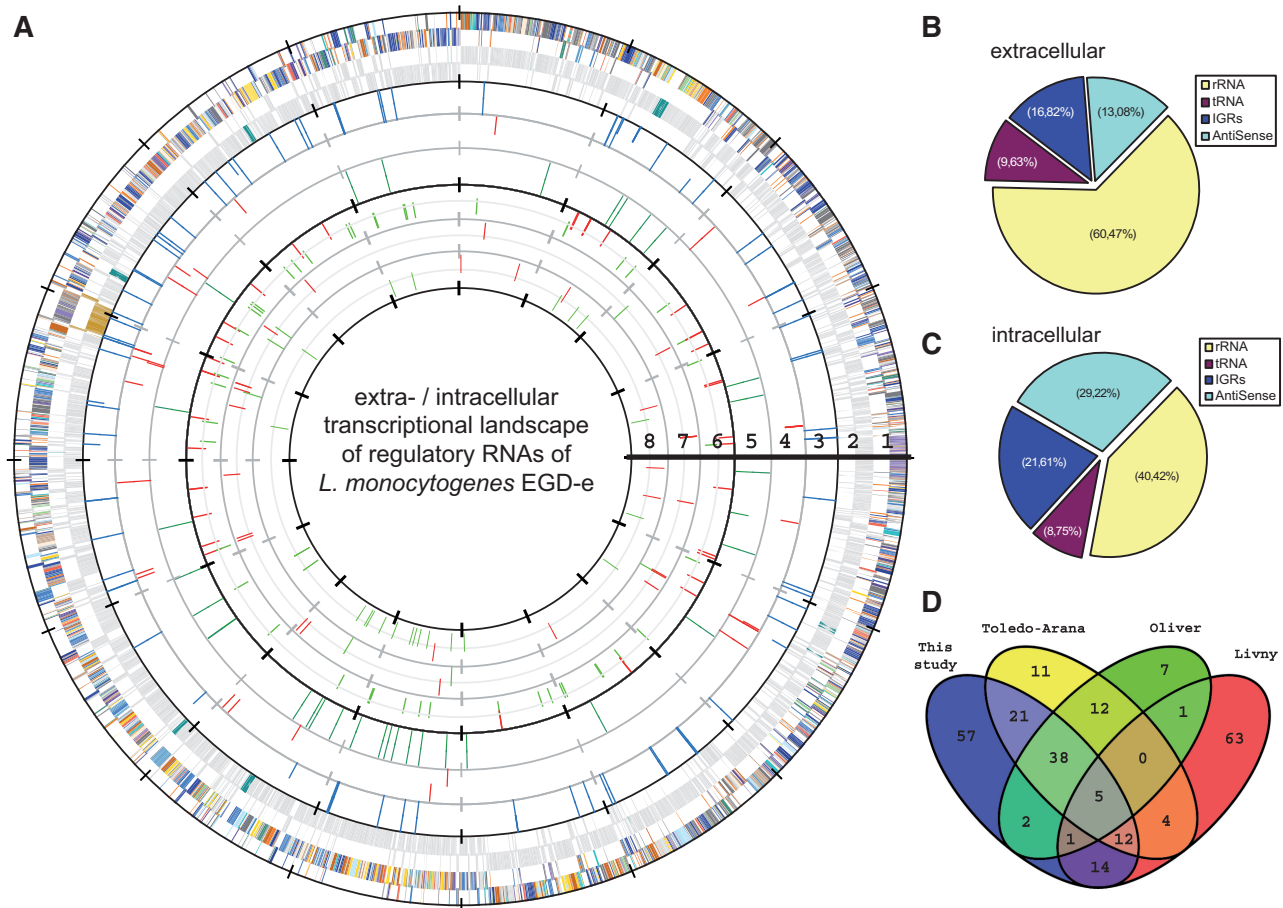
Figure 1. Discovery of the intracellular sRNome of *L. monocytogenes* using RNA-Seq. (**A**) Extracellular and intracellular transcriptional landscape of *L. monocytogenes* is represented using GenomeViz (46). Circles display following information from outside to inside: (1) COG categories; (2) rRNAs and tRNAs (blue), a prophage-like locus (light brown) and the virulence gene cluster (red); (3) intracellular regulatory RNAs (outer circle) and extracellular regulatory RNAs (inner circle); (4) intracellular asRNAs (outer circle) and extracellular asRNAs (inner circle); (5) intracellular *cis*-regulatory RNAs including riboswitches (outer circle) and extracellular *cis*-regulatory RNAs including riboswitches (inner circle); (6) regulation of intracellular sRNAs; (7) regulation of intracellular asRNAs and (8) intracellular *cis*-regulatory RNAs including riboswitches; (**B** and **C**) Distribution of mapped sequence reads used for extracellular and intracellular transcriptome analysis; (**D**) Comparative analysis of sRNA transcriptome data using 'cumulative' values which can be summarized since sRNA candidates would not be counted multiple times (see Supplementary Figure S1 for a 'non-cumulative' version). Comparison of our RNA-seq results, whole genome tiling array from Toledo-Arana and coworkers (31), RNA-seq data of *L. monocytogenes* 10403S (32) and *in silico* regulatory RNA predictions (1).

The vast majority of the putative promoters were accounted for by the sigma70 promoter which was identified preceding 28 intracellular and 34 extracellular regulatory RNAs, respectively. For six intracellular and four extracellular candidates a SigB box was detected in the upstream region suggesting an involvement in the general stress response. The PrfA regulator which is involved in virulence (51,52), putatively binds to the upstream region of two intracellular and one extracellular candidates. We used a pre-computed rho-independent transcription terminator prediction from TransTermHP (41) to identify transcriptional units. For 55% of the sRNAs we were able to define a putative terminator (Supplementary Table S3).

Chromosomal distribution of regulatory RNA

We used GenomeViz (46) to visualize the location and distribution of the regulatory sRNAs on the *L. monocytogenes* EGD-e genome. We found that they

are neither located within specific chromosomal clusters nor do they exhibit a specific strand prevalence (Figure 1A).

Comparative analysis with whole genome tiling array data, RNA-seq and *in silico* predictions

We compared 103 putative sRNAs previously published by Toledo-Arana and colleagues to our 150 identified regulatory RNA candidates (53). The comparison revealed 75% of previously published sRNA candidates by Toledo-Arana (31) to be overlapping with the candidates of the current study (Figure 1D and Supplementary Table S3). Of 26 candidates identified by Toledo-Arana but not identified in our study no expression was observed for 12 candidates, 12 regulatory RNAs failed to reach predefined cutoffs and two were either expressed on the opposite strand or were found in the same IGR but without any overlap.

Comparison of our data to that of Toledo-Arana (31) from bacteria grown in BHI, revealed an overlap of 68% of all sRNAs. A comparison of sRNAs expressed when grown in blood by Toledo-Arana to the intracellular sRNA repertoire detected in our study showed considerable overlap of detected regulatory RNAs (72%). In some cases such as for sRNA candidate *rli33*, which is 534 nt in size and represents the longest sRNA in the published work (31), our analysis yielded two different sRNA candidates. In order to avoid confusion we designate them *rli33-1* and *rli33-2*.

In addition we compared Illumina deep sequencing results previously reported by Oliver *et al.* (32) and detected an overlap of 69% with our work (Figure 1D and Supplementary Table S3). Computational analysis using the SIPHT workflow has previously predicted 100 regulatory RNAs for *L. monocytogenes* EGD-e (1). Of these, only 33% could be recovered by our experiment (Figure 1D and Supplementary Table S3) whereas the remaining 77 predicted candidates show no expression or failed to attain the desired cutoffs.

For 14 regulatory RNA candidates we predicted small ORFs (Supplementary Table S3) applying a gene calling cutoff more than 10 amino acids, which may not be represented in the current version of the *L. monocytogenes* genome annotation (36).

We reanalyzed the *rliB* locus which is involved in virulence (31) using the RNA-seq data described herein. This locus harbors five small copies of expressed CRISPR repeats (29 nt long) and spacers (36 nt long) sequences (crRNA) predicted by the CRISPR prediction software PILER-CR (54) and could be confirmed by data present in this study and which was originally noticed by Mandin *et al.* (29). However, no homologous of CRISPR-associated (CAS) genes were detected in the flanking regions. The predicted CRISPR spacer and repeats have a range of between 24–47 bp and 26–72 bp, respectively, and are currently the smallest predicted sRNAs for *L. monocytogenes*.

Verification of selected sRNA candidates

To verify the transcription of sRNA candidates at the locations indicated by deep sequencing, we used northern blot analysis. Four sRNA candidates namely *rli31*, *rli33-1, rli50* and *rli112* were selected for verification according to their high read numbers revealed by sequencing results.

The results of the northern blotting, shown in Figure 2, demonstrate signals corresponding to small transcripts from each of the candidates. According to the signals in the northern blot analysis, some of the sRNA candidates were differentially expressed under the tested conditions. *rli31* and *rli33-1* showed variation in transcript concentration during extra- and intra-cellular growth. Both sRNA candidates exhibited significantly higher transcript numbers during intracellular growth. The northern blot results correlate with read-numbers from transcriptome deep sequencing (Supplementary Table S3). The determined size of *rli33-1* by northern blot analysis of ~140 nt confirms the read length (144 nt) assessed by
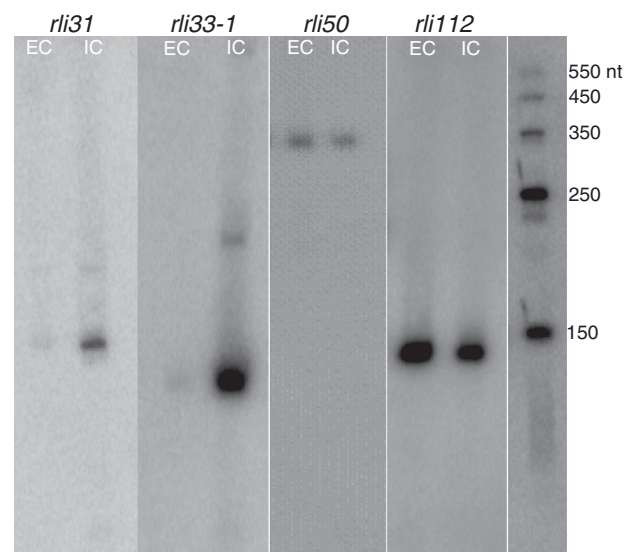
Figure 2. Northern blots of sRNA candidates. Validation of RNA-seq data with northern blot analysis of extracellularly expressed sRNAs from bacteria at mid exponential growth phase in BHI (EC) compared with intracellularly expressed sRNAs at 4 h post-infection in murine macrophages (IC) of *rli31*, *rli33-1*, *rli50* and *rli112*.

RNA-seq. *rli33-1* showed two bands in the northern blot at ~200 and 130 nt in length, that probably originate from processing of the larger co-transcript. Both fragments were also detected by transcriptome sequencing excluding the possibility of cross-hybridization (Figure 2). *rli50* appears to be longer than predicted by sequencing (306 nt) and has a size of ~350 nt as determined by northern blot. We observed with an *rli112* specific probe strong northern blot signals (~140 nt) under extra- and intra-cellular conditions indicating high RNA abundance of this new unpublished putative sRNA.

Noteworthy is the observation that a homolog of *rli112*, *rli78*, is highly expressed under extracellular growth conditions compared with intracellular growth. To exclude effects of cross hybridization in the northern blot analysis, qRT–PCR was conducted with isogenic mutants lacking *rli78* and *rli112*, respectively, using a primer pair that binds to the homologous region in the same location. The results confirmed higher expression levels of *rli78* under extracellular growth conditions (data not shown). This result correlates with the transcript numbers determined by RNA sequencing and reflect higher cDNA reads extracellularly as indicated by sequencing data (Supplementary Table S3). Additionally, six sRNA candidates including three so far new unpublished sRNAs (*rli80*, *rli91* and *rli105*) were confirmed (Supplementary Figure S2). Generally, northern blotting confirmed the existence of these sRNA candidates and showed that sizes, except *rli91*, correlate with those predicted by RNA-seq. Quantitation based on sequencing frequently differs significantly from that observed on northern blots which is exemplified by *rli80* and *rli105* (Supplementary Figure S2). Further investigations are required to investigate the basis of these differences.
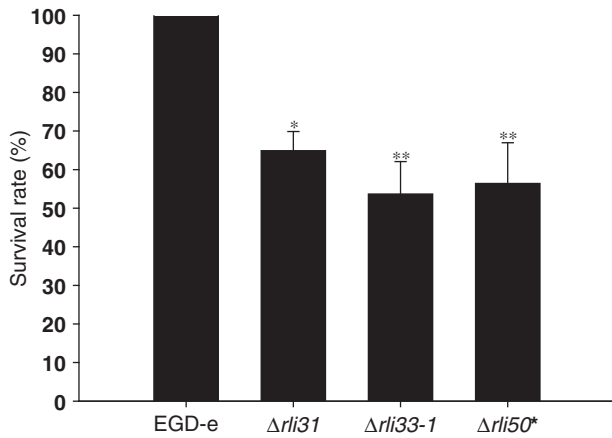
Figure 3. Survival of *L. monocytogenes* sRNA mutants in P388D1 murine macrophage cells. The macrophages were infected with the wild-type *L. monocytogenes* EGD-e and its isogenic deletion mutants, *rli31*, *rli33-1* and *rli50**, with an MOI of 10 in 24-well plates and bacterial CFU counts were measured on agar plated following lysis of the P388D1 cells after 4 h post-infection. *n* = 5; error bars indicate standard deviations (*$P \leq 0.005$, **$P \leq 0.05$).

Moreover, our study provides the first experimental evidence for the presence of two antisense RNAs, *anti2394* and *anti2095*, in *L. monocytogenes* EGD-e (Supplementary Figure S3).

Mutants lacking *rli31*, *rli33-1* and *rli50** are attenuated for infection

Due to an overlap between *rli50* (minus strand) and *rli112* (plus strand) of ~43% in the IGR between *lmo2709* and *lmo2710* we decided to delete the distal end of *rli112* to create the mutant Δ*rli50** (Δ*rli50** refers to the deletion mutant and *rli50* to the sRNA). We generated isogenic deletion mutants of the sRNA candidates *rli31*, *rli33-1* and *rli50** with the highest intracellular transcription levels (Supplementary Figures S4 and S5). Additionally, these sRNA candidates had higher intracellular cDNA read numbers in comparison to extracellular growth conditions. A schematic overview of Δ*rli50** and Δ*rli33-1* and is presented in Supplementary Figure S4.

We assessed the mutant strains' capability to grow in P388D1 murine macrophages together with the wild-type strain. Whereas the mutants Δ*rli31*, Δ*rli33-1* and Δ*rli50** were significantly impaired in their abilities to proliferate intracellularly in macrophages as compared with the wild-type strain (Figure 3), no differences were observable in their *in vitro* growth in BHI (Supplementary Figure S6).

We investigated the survival of larvae from *Galleria* following injection with different sRNA mutants of *L. monocytogenes*. Bacterial cultures grown to exponential phase were injected dorsolaterally into the hemocoel at 10^6 CFU/larva. We observed significant attenuation in the mortality rates of larvae when injected with Δ*rli31*, Δ*rli33-1* and Δ*rli50** in comparison to wild-type EGD-e (Figure 4A–C).

We used the mouse infection model to assess the virulence properties of the sRNA mutants Δ*rli31*, Δ*rli33-1*
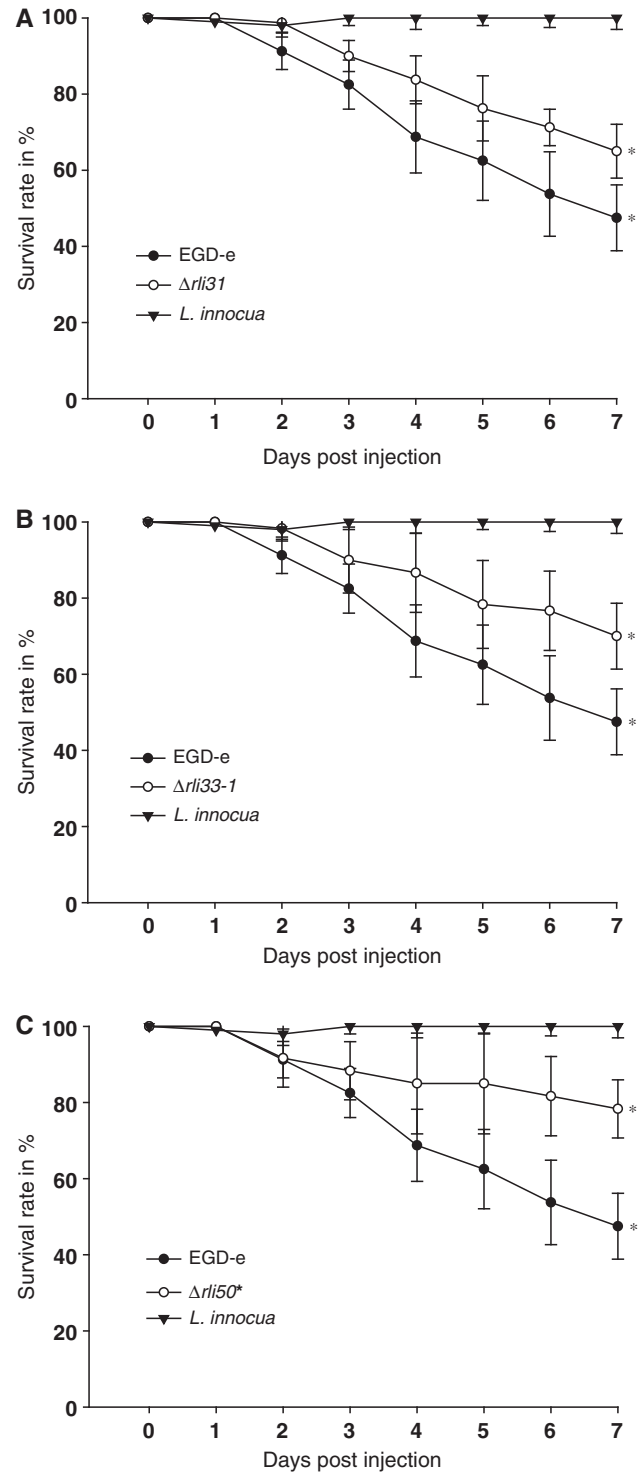
Figure 4. Survival of *Galleria mellonella* larvae after inoculation with different *L. monocytogenes* sRNA mutants and *L. innocua*. Time course of survival of the larvae varies with the type of sRNA mutants employed for inoculation. Inoculation with 10^6 CFU/larvae EGD-e resulted in significantly higher killing rate of larvae in comparison to (**A**) *rli31*, (**B**) *rli33-1* and (**C**) *rli50**. The non-pathogenic *L. innocua* showed no mortality. Values represent means of at least three independent experiments ± standard deviations for 20 larvae per treatment (*$P \leq 0.005$).
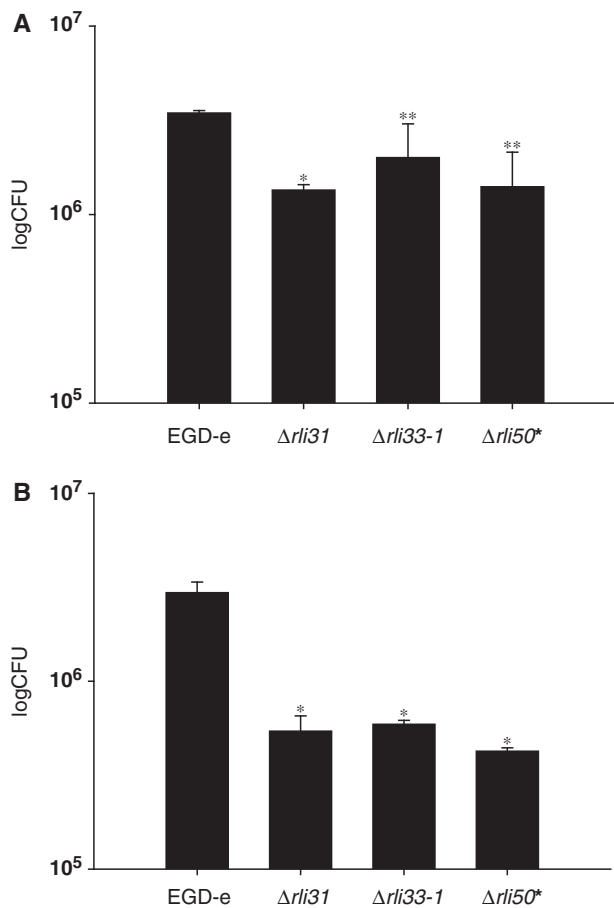
Figure 5. Mice infection studies with sRNA deletion mutants and of *L. monocytogenes*. Bacterial load in mice organs were also determined following *in vitro* infection with 2000 CFU of *L. monocytogenes* EGD-e wild-type strain as well as its isogenic sRNA mutants *rli31*, *rli33-1* and *rli50**. On day 3 after infection, the numbers of viable bacteria in spleens (**A**) and livers (**B**) of three animals per group were determined of wild-type EGD-e versus *rli31*, *rli33-1* and *rli50** in spleen and liver, respectively ($n = 4$). Error bars indicate standard deviations ($*P \leq 0.005$, $**P < 0.05$).

and $\Delta rli50*$ to that of wild-type *L. monocytogenes* EGD-e. For all three mutants, survival and/or growth in the spleen, and in particular in the liver, were significantly reduced at day 3 post-infection when compared with wild-type (Figure 5A and B).

To exclude that the phenotypes of the deletion mutants were due to polar effects on flanking genes, we generated complemented strains of the $\Delta rli31$, $\Delta rli33-1$ and $\Delta rli50*$ sRNA deletion mutants and examined their abilities to proliferate intracellularly in P388D1 murine macrophage cells. The complementation largely restored the intracellular growth impairment in macrophages (Supplementary Figure S7, see also Figure 3). Additionally, we checked the expression of the flanking genes by quantitative real time PCR under both conditions used in the experiment (extracellular versus intracellular). The results presented in Supplementary Figure S8 show clearly that the deletion of the putative sRNA candidates has no consequences on the expression of the flanking genes.

Comparative analysis of putative regulatory RNAs among members of the genus *Listeria*

Finally, we compared our RNA sequencing identified regulatory RNA to the genomes of four human pathogenic strains of *L. monocytogenes* (36,53,54), an attenuated *L. monocytogenes* 4a strain (HCC23) (56) and the three apathogenic species of *L. innocua*, *L. welshimeri* and *L. seeligeri* (36,43,44) to investigate the regulation of virulence. Here we detected that the number of regulatory RNAs among the *L. monocytogenes* serotypes 4b and 1/2a belonging to Lineage I and II were highly similar compared with *L. monocytogenes* 1/2a EGD-e as reference (II > I > III), but decreased rapidly, when listerial species were more distantly related (Figure 6).

DISCUSSION

In this study, we identified that a total of 150 putative regulatory RNAs were expressed during growth of *L. monocytogenes* in extracellular and intracellular environments. The putative regulatory RNAs can be divided into three groups including 88 putative sRNAs, 29 asRNAs and 33 *cis*-regulatory elements including riboswitches which are expressed during growth of *L. monocytogenes* in extracellular and intracellular environments.

Previously, Toledo-Arana used genome-wide tiling arrays and reported a total 103 sRNAs expressed by *L. monocytogenes* growing under a wide range of conditions from broth culture to blood and the intestinal lumen (31). In this study we detected 79 of the regulatory RNAs previously described suggesting good correlation of the data, given the inherent differences in comparing sequencing-based technologies with that of hybridization analyses. Nevertheless, even though we have only compared two growth conditions using deep sequencing of cDNA derived from size-fractionated RNA (<500 nt), our data revealed 71 new candidates, none of which have been previously described. Our findings suggest that diversity of regulation at the post-transcriptional level is an important component of adaptation to niche specific growth.

For functional analysis we focused on the intracellularly up-regulated sRNA candidates $\Delta rli31$, $\Delta rli33-1$ and $\Delta rli50*$. Isogenic deletion mutants of these sRNA loci resulted in reproducibly reduced growth properties following infection of P388D1 murine macrophages and in virulence attenuation in both insect and mice models of infection. Because *G. mellonella* is a model for innate immunity responses, it is likely that *rli31*, *rli33-1* and *rli50** represent sRNA loci required for adaptation to intracellular growth in vertebrate and invertebrate hosts.

The gene *lmo0559* located downstream of *rli31* encodes a putative transporter of the CorA superfamily involved in magnesium and cobalt uptake. Intracellular transcriptomic analysis of *L. monocytogenes* has previously indicated that the host cytosol is a rich source of ions as suggested by the down-regulation of several ion transport systems (37). An involvement of the 5′-UTR in the regulation of *mgtA*, a magnesium transporter, has
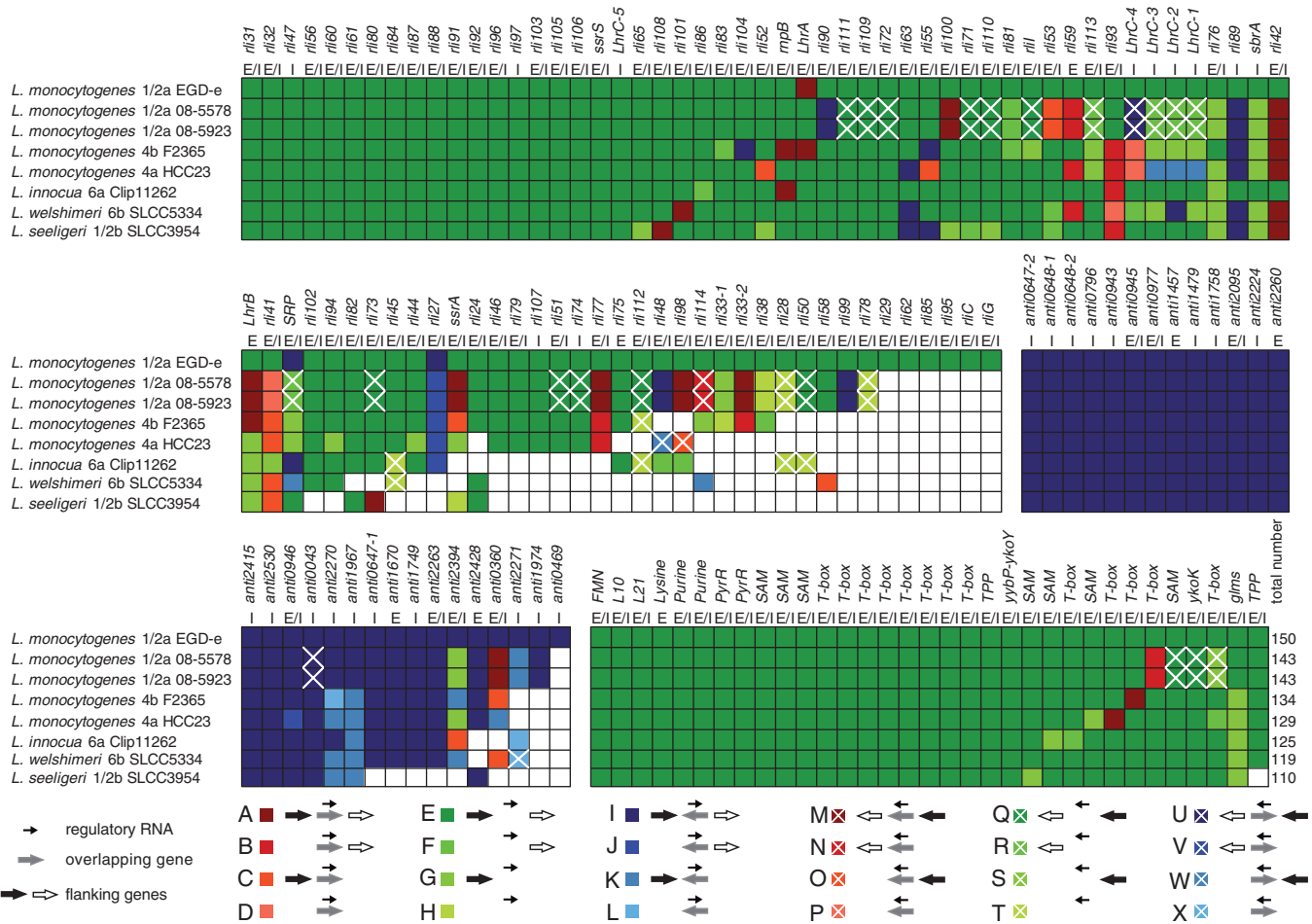
Figure 6. Comparative overview of known and putative regulatory RNAs of *L. monocytogenes* EGD-e. EGD-e was compared with 3 *L. monocytogenes* serotypes (3 × 1/2a, 1 × 4b and 1 × 4a) and three non-pathogenic *Listeria* species (*L. innocua*, *L. welshimeri* and *L. seeligeri*). To determine the distribution of regulatory RNA inside the genus a BLAST analysis was conducted using sRNAdb (unpublished software). Candidates were considered present inside a strain in the case of a sequence identity of 60% and a coverage of 80%. Since the surrounding locus is often important for the function of the regulatory RNA, information about the conservation of adjacent genes was included using the same cutoff. Possible cases for direction, presence and absence of each regulatory RNA and its flanking genes was color-coded below. A white square indicates the absence of the regulatory RNA. As a reference for this analysis the relevant loci of *L. monocytogenes* EGD-e were chosen. The small black arrow depicted in the legend indicates the regulatory RNA, while larger arrows in black and white symbolize the left and right flanking gene, respectively. A large gray arrow denotes a gene overlapping the regulatory RNA in sense or antisense direction. The arrow direction is not representative for the strand but for the relation to the locus in the reference genome of *L. monocytogenes* EGD-e. (E) indicates extracellular and (I) intracellular expression of the regulatory RNA.

been reported previously in *Salmonella enterica* (57), where high Mg^{2+} concentration induced the formation of a stem–loop structure leading to the transcriptional termination of *mgtA*.

Interestingly, a putative sRNA located in the intergenic region between *lmo0671* and *lmo0672* was previously reported as *rli33* encoding a transcript of 534 nt in length (31). However, our RNA-seq revealed two smaller fragments of 186 nt (*rli33-1*) and 274 nt (*rli33-2*) in this region. In support of this observation, northern blot analyses confirmed the presence of *rli33-1* and *rli33-2* suggesting either the presence of an internal start site within the larger transcript or an unknown RNA processing mechanism of the transcript under the experimental conditions used. *rli33* is highly induced in the stationary growth phase and in blood indicating a potential role in virulence (31). This has been shown by *in vitro*

and *in vivo* experiments with the isogenic deletion mutant *rli33-1* which completely removes the 5′-end of *rli33*.

The *rli50** deletion mutant showed the strongest effect in the insect model and in the liver of mice. Interestingly, our RNA-seq data revealed two sRNA candidates in the chromosomal region between *lmo2709* and *lmo2710*, a new one on the sense (*rli112*) and the other on the antisense strand (*rli50*), which was detected by Toledo-Arana and coworkers (31). As mentioned above the isogenic mutant for *rli50**, removes the distal end of the *rli112* sRNA candidate resulting in a double mutant that may be responsible for the impaired virulence effects obtained. A homolog of *rli112* with 94% similarity is located within the intergenic region between the genes *lmo0470* and *lmo0471*, designated as sRNA candidate *rli78*, which has higher extracellular expression than intracellularly. However, this sRNA is clearly not able to compensate

for the loss of sRNA *rli50* in the isogenic mutant of *rli50**. Northern blot analysis of *rli29* and *rli78* confirmed the presence of the sense and antisense strand transcripts (Supplementary Figure S2). This indicates different regulatory roles of these homologous sRNAs in their environmental niche.

Similar examples for the existence of several copies of sRNA were reported for *V. cholerae,* which harbours four redundant copies of the regulatory RNA Qrr (Qrr1-4) interacting with two feedback loops and which promotes gene dosage compensation among Qrr1-4 for proper control of quorum-sensing processes (58,59). Also, multiple copies of other listerial sRNAs are currently known for the Hfq-dependent LhrC (28) with five copies in total and present at two different chromosomal locations. We find that LhrC is up-regulated in our intracellular data, but the role of these intracellularly induced sRNAs has still to be elucidated. Surprisingly, the chromosomal locus of *lmo0459–lmo0479,* including *rli28, rli29* and *rli78*, revealed a lower GC-content suggesting horizontal gene transfer events. An inspection of the surrounding loci indicated a gene encoding a transposase of the IS3 family (*lmo0464*) in this region. Thus, we speculate that horizontal gene transfer (HGT) caused by transposition might be involved in chromosomal spreading of listerial regulatory RNAs.

Our data also indicate differential regulation of sRNA candidates within the A118 prophage-like regions (*rliG*, *rli48*, *rli62*, *rli98* and *rli99*). Bacteriophages play an important role for horizontal gene transfer to shape their microbial host genomes with new genetic functions including sRNA, such as has been described for *ipeX* which is responsible for OmpC porin regulation in *E. coli* (60). Transcriptional activation of phage-related genes has been previously reported from different groups by *in vivo* transcriptomic studies (37,61,62). Further investigation is required to understand the contribution of these phage-related sRNA transcripts to the infection process of *L. monocytogenes.*

In addition, CRISPR systems, responsible for microbial phage defense, are elements that have been described in a large number of prokaryotic genomes (63). We have confirmed as originally noticed by Mandin *et al*. (29) that *rliB* is related to CRISPR elements, but CRISPR-associated CAS genes were absent suggesting that these transcripts are produced by a novel endonucleolytic mechanism.

An emerging class of regulatory RNAs are antisense RNAs which have been observed in different bacteria and archaea (17,64–67). This class of molecules was previously known only for transposons, plasmids and phages (68–70). Transcriptome analysis of *E. coli* and *H. pylori* indicated antisense transcription across the entire genome (17,71). In addition, several long and short antisense RNAs were previously reported for *Listeria* (31). We have identified a number of novel short antisense RNAs in our experiments (Supplementary Table S3). The vast majority of these were intracellularly up-regulated, although there was no strict correlation to down-regulation of their potential targeted genes when inspecting microarray-based intracellular transcriptome studies (37,61). Studies of individual plasmid-encoded and chromosomally encoded asRNAs in a variety of bacterial species have demonstrated that asRNAs can regulate gene expression at the level of translation, mRNA stability or transcription (68).

We were able to demonstrate, for the first time, the existence of two asRNAs (*anti2095* and *anti2394*) with increased intracellular expression using northern blots. Both asRNAs were previously reported to have short and long variants (31). Due to RNA-seq using RNA fraction <500 nt, only the short asRNAs were confirmed in our study (Supplementary Figure S3). The genes transcribed from the opposite strand encoding a hypothetical protein (*lmo2394*) and a phosphofructokinase (*lmo2095*) did not show significant expression changes after shifting from extra- to intra-cellular conditions (37,61). Also a genome-wide COG analysis of asRNA–mRNA targets revealed no obvious correlation to specific classes of genes suggesting that asRNAs might be used to control general processes within the bacterial cell.

Riboswitches are *cis*-acting RNA structures responsible for downstream regulation of gene expression in bacteria (15). In *B. subtillis* 2% of the genes are regulated by these RNA elements (72). Out of 42 known listerial riboswitches in the RFAM database, 33 were detected in our study and the majority of them were down-regulated intracellularly. Interestingly, the lysine riboswitch (31) was not expressed during intracellular survival of bacteria indicating a lysine restricted intracellular environment for the bacterial pathogen.

Loh and colleagues have reported on the ability of two S-adenosylmethionine (SAM) riboswitches, SreA and SreB, to act *in trans* to modulate the expression of the critical listerial virulence regulator PrfA (73). We have observed that the SreA and SreB are down-regulated within the host cytosol indicating that the decreased copy number of SreA and SreB allow an intracellular induction of the PrfA regulator. The T-box class of riboswitches represents the largest class of riboswitches in *L. monocytogenes* sensing the level of uncharged tRNAs in the bacterial cell, which is induced in our data. The expression of the corresponding tRNA synthetase genes have previously been found to be decreased in the host cytosolic phase (37), which suggests the involvement of T-box regulation due to infection.

Finally, we found that listerial species have highly conserved riboswitches for adaption to physiological processes, but have clear differences in their sRNA and asRNA repertoire suggesting adaptation to their potential ecological habitats (Figure 6). Indeed comparative analysis indicated several strain and serotype-specific putative regulatory RNAs such as *rli112* which enables this sRNA to become a potential diagnostic marker for lineage II.

CONCLUSION

Here we show that extensive expression of sRNA candidates in *L. monocytogenes* occurs during intracellular growth. Our studies uncovered 71 previously undescribed

putative regulatory RNAs and revealed 29 candidates that are specifically expressed during intracellular growth. Although we have shown that several of these sRNAs are indeed required for virulence of the bacterium their precise role in promoting bacterial survival under these conditions remains to be studied. As a next step, the identification of targets of these putative regulatory RNAs will be required and will enable us to understand the regulatory response triggered by the bacterium when shifting from extracellular to intracellular growth conditions.

ACCESSION NUMBERS

RNA sequencing data have been deposited to ArrayExpress (http://www.ebi.ac.uk/ arrayexpress), accession number E-MTAB-329.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Livny,J., Teonadi,H., Livny,M. and Waldor,M.K. (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS ONE*, **3**, e3197.
2. Storz,G., Altuvia,S. and Wassarman,K.M. (2005) An abundance of RNA regulators. *Annu. Rev. Biochem.*, **74**, 199–217.
3. Vogel,J. and Sharma,C.M. (2005) How to find small non-coding RNAs in bacteria. *Biol. Chem.*, **386**, 1219–1238.
4. Gottesman,S. (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet.*, **21**, 399–404.
5. Sorek,R. and Cossart,P. (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.*, **11**, 9–16.
6. Livny,J. and Waldor,M.K. (2010) Mining regulatory 5′UTRs from cDNA deep sequencing datasets. *Nucleic Acids Res.*, **38**, 1504–1514.
7. Vogel,J., Bartels,V., Tang,T.H., Churakov,G., Slagter-Jager,J.G., Huttenhofer,A. and Wagner,E.G. (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.*, **31**, 6435–6443.
8. Washietl,S., Hofacker,I.L., Lukasser,M., Huttenhofer,A. and Stadler,P.F. (2005) Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nat. Biotechnol.*, **23**, 1383–1390.
9. Washietl,S., Hofacker,I.L. and Stadler,P.F. (2005) Fast and reliable prediction of noncoding RNAs. *Proc. Natl Acad. Sci. USA*, **102**, 2454–2459.
10. Zhang,A., Wassarman,K.M., Rosenow,C., Tjaden,B.C., Storz,G. and Gottesman,S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.*, **50**, 1111–1124.
11. Mraheil,M.A., Billion,A., Kuenne,C., Pischimarov,J., Kreikemeyer,B., Engelmann,S., Hartke,A., Giard,J.C., Rupnik,M., Vorwerk,S. *et al.* (2010) Comparative genome-wide analysis of small RNAs of major Gram-positive pathogens: from identification to application. *Microb. Biotechnol.*, **3**, 658–676.
12. Marker,C., Zemann,A., Terhorst,T., Kiefmann,M., Kastenmayer,J.P., Green,P., Bachellerie,J.P., Brosius,J. and Huttenhofer,A. (2002) Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. *Curr. Biol.*, **12**, 2002–2013.
13. Gottesman,S. (2004) The small RNA regulators of *Escherichia coli*: roles and mechanisms*. *Annu. Rev. Microbiol.*, **58**, 303–328.
14. Repoila,F. and Darfeuille,F. (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol. Cell*, **101**, 117–131.
15. Waters,L.S. and Storz,G. (2009) Regulatory RNAs in bacteria. *Cell*, **136**, 615–628.
16. Romby,P., Vandenesch,F. and Wagner,E.G. (2006) The role of RNAs in the regulation of virulence-gene expression. *Curr. Opin. Microbiol.*, **9**, 229–236.
17. Sharma,C.M., Hoffmann,S., Darfeuille,F., Reignier,J., Findeiss,S., Sittka,A., Chabas,S., Reiche,K., Hackermuller,J., Reinhardt,R. *et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, **464**, 250–255.
18. Toledo-Arana,A., Repoila,F. and Cossart,P. (2007) Small noncoding RNAs controlling pathogenesis. *Curr. Opin. Microbiol.*, **10**, 182–188.
19. Franze de Fernandez,M.T., Eoyang,L. and August,J.T. (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature*, **219**, 588–590.
20. Franze de Fernandez,M.T., Hayward,W.S. and August,J.T. (1972) Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. *J. Biol. Chem.*, **247**, 824–831.
21. Valentin-Hansen,P., Eriksen,M. and Udesen,C. (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol. Microbiol.*, **51**, 1525–1533.
22. Aiba,H. (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.*, **10**, 134–139.
23. Ding,Y., Davis,B.M. and Waldor,M.K. (2004) Hfq is essential for *Vibrio cholerae* virulence and downregulates sigma expression. *Mol. Microbiol.*, **53**, 345–354.
24. McNealy,T.L., Forsbach-Birk,V., Shi,C. and Marre,R. (2005) The Hfq homolog in *Legionella pneumophila* demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. *J. Bacteriol.*, **187**, 1527–1532.
25. Roop,R.M., Robertson,G.T., Ferguson,G.P., Milford,L.E., Winkler,M.E. and Walker,G.C. (2002) Seeking a niche: putative contributions of the *hfq* and *bacA* gene products to the successful adaptation of the brucellae to their intracellular home. *Vet. Microbiol.*, **90**, 349–363.
26. Sonnleitner,E., Hagens,S., Rosenau,F., Wilhelm,S., Habel,A., Jager,K.E. and Blasi,U. (2003) Reduced virulence of a *hfq* mutant of *Pseudomonas aeruginosa* O1. *Microb. Pathog.*, **35**, 217–228.
27. Christiansen,J.K., Larsen,M.H., Ingmer,H., Sogaard-Andersen,L. and Kallipolitis,B.H. (2004) The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. *J. Bacteriol.*, **186**, 3355–3362.
28. Christiansen,J.K., Nielsen,J.S., Ebersbach,T., Valentin-Hansen,P., Sogaard-Andersen,L. and Kallipolitis,B.H. (2006) Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. *RNA*, **12**, 1383–1396.
29. Mandin,P., Repoila,F., Vergassola,M., Geissmann,T. and Cossart,P. (2007) Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res.*, **35**, 962–974.
30. Nielsen,J.S., Olsen,A.S., Bonde,M., Valentin-Hansen,P. and Kallipolitis,B.H. (2008) Identification of a sigma B-dependent

small noncoding RNA in *Listeria monocytogenes*. *J. Bacteriol.*, **190**, 6264–6270.

31. Toledo-Arana,A., Dussurget,O., Nikitas,G., Sesto,N., Guet-Revillet,H., Balestrino,D., Loh,E., Gripenland,J., Tiensuu,T., Vaitkevicius,K. *et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature*, **459**, 950–956.

32. Oliver,H.F., Orsi,R.H., Ponnala,L., Keich,U., Wang,W., Sun,Q., Cartinhour,S.W., Filiatrault,M.J., Wiedmann,M. and Boor,K.J. (2009) Deep RNA sequencing of *L. monocytogenes* reveals overlapping and extensive stationary phase and sigma B-dependent transcriptomes, including multiple highly transcribed noncoding RNAs. *BMC Genomics*, **10**, 641.

33. Duffy,L.L., Vanderlinde,P.B. and Grau,F.H. (1994) Growth of *Listeria monocytogenes* on vacuum-packed cooked meats: effects of pH, aw, nitrite and ascorbate. *Int. J. Food Microbiol.*, **23**, 377–390.

34. Cossart,P. and Toledo-Arana,A. (2008) *Listeria monocytogenes*, a unique model in infection biology: an overview. *Microbes. Infect.*, **10**, 1041–1050.

35. Hain,T., Chatterjee,S.S., Ghai,R., Kuenne,C.T., Billion,A., Steinweg,C., Domann,E., Karst,U., Jansch,L., Wehland,J. *et al.* (2007) Pathogenomics of *Listeria* spp. *Int. J. Med. Microbiol.*, **297**, 541–557.

36. Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A., Baquero,F., Berche,P., Bloecker,H., Brandt,P., Chakraborty,T. *et al.* (2001) Comparative genomics of *Listeria* species. *Science*, **294**, 849–852.

37. Chatterjee,S.S., Hossain,H., Otten,S., Kuenne,C., Kuchmina,K., Machata,S., Domann,E., Chakraborty,T. and Hain,T. (2006) Intracellular gene expression profile of *Listeria monocytogenes*. *Infect. Immun.*, **74**, 1323–1338.

38. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.

39. Griffiths-Jones,S., Moxon,S., Marshall,M., Khanna,A., Eddy,S.R. and Bateman,A. (2005) Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.*, **33**, D121–D124.

40. Eddy,S.R. (2009) A new generation of homology search tools based on probabilistic inference. *Genome Inform.*, **23**, 205–211.

41. Kingsford,C.L., Ayanbule,K. and Salzberg,S.L. (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.*, **8**, R22.

42. Meyer,F., Goesmann,A., McHardy,A.C., Bartels,D., Bekel,T., Clausen,J., Kalinowski,J., Linke,B., Rupp,O., Giegerich,R. *et al.* (2003) GenDB–an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.*, **31**, 2187–2195.

43. Hain,T., Steinweg,C., Kuenne,C.T., Billion,A., Ghai,R., Chatterjee,S.S., Domann,E., Karst,U., Goesmann,A., Bekel,T. *et al.* (2006) Whole-genome sequence of *Listeria welshimeri* reveals common steps in genome reduction with *Listeria innocua* as compared to *Listeria monocytogenes*. *J. Bacteriol.*, **188**, 7405–7415.

44. Steinweg,C., Kuenne,C.T., Billion,A., Mraheil,M.A., Domann,E., Ghai,R., Barbuddhe,S.B., Karst,U., Goesmann,A., Puhler,A. *et al.* (2010) Complete genome sequence of Listeria seeligeri, a nonpathogenic member of the genus *Listeria*. *J. Bacteriol.*, **192**, 1473–1474.

45. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics.*, **25**, 2078–2079.

46. Ghai,R., Hain,T. and Chakraborty,T. (2004) GenomeViz: visualizing microbial genomes. *BMC Bioinformatics.*, **5**, 198.

47. Thedieck,K., Hain,T., Mohamed,W., Tindall,B.J., Nimtz,M., Chakraborty,T., Wehland,J. and Jansch,L. (2006) The MprF protein is required for lysinylation of phospholipids in listerial membranes and confers resistance to cationic antimicrobial peptides (CAMPs) on *Listeria monocytogenes*. *Mol. Microbiol.*, **62**, 1325–1339.

48. Schaferkordt,S. and Chakraborty,T. (1995) Vector plasmid for insertional mutagenesis and directional cloning in *Listeria* spp. *Biotechniques*, **19**, 720–725.

49. Lauer,P., Chow,M.Y., Loessner,M.J., Portnoy,D.A. and Calendar,R. (2002) Construction, characterization, and use of two *Listeria monocytogenes* site-specific phage integration vectors. *J. Bacteriol.*, **184**, 4177–4186.

50. Mukherjee,K., Altincicek,B., Hain,T., Domann,E., Vilcinskas,A. and Chakraborty,T. (2010) *Galleria mellonella* as a model system for studying *Listeria* pathogenesis. *Appl. Environ. Microbiol.*, **76**, 310–317.

51. Chakraborty,T., Leimeister-Wachter,M., Domann,E., Hartl,M., Goebel,W., Nichterlein,T. and Notermans,S. (1992) Coordinate regulation of virulence genes in *Listeria monocytogenes* requires the product of the *prfA* gene. *J. Bacteriol.*, **174**, 568–574.

52. Hamon,M., Bierne,H. and Cossart,P. (2006) *Listeria monocytogenes*: a multifaceted model. *Nat. Rev. Microbiol.*, **4**, 423–434.

53. Nelson,K.E., Fouts,D.E., Mongodin,E.F., Ravel,J., DeBoy,R.T., Kolonay,J.F., Rasko,D.A., Angiuoli,S.V., Gill,S.R., Paulsen,I.T. *et al.* (2004) Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res.*, **32**, 2386–2395.

54. Edgar,R.C. (2007) PILER-CR: fast and accurate identification of CRISPR repeats. *BMC Bioinformatics*, **8**, 18.

55. Gilmour,M.W., Graham,M., Van Domselaar,G., Tyler,S., Kent,H., Trout-Yakel,K.M., Larios,O., Allen,V., Lee,B. and Nadon,C. (2010) High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. *BMC Genomics*, **11**, 120.

56. Liu,D., Lawrence,M.L., Pinchuk,L.M., Ainsworth,A.J. and Austin,F.W. (2007) Characteristics of cell-mediated, anti-listerial immunity induced by a naturally avirulent *Listeria monocytogenes* serotype 4a strain HCC23. *Arch. Microbiol.*, **188**, 251–256.

57. Cromie,M.J., Shi,Y., Latifi,T. and Groisman,E.A. (2006) An RNA sensor for intracellular Mg(2+). *Cell*, **125**, 71–84.

58. Lenz,D.H., Mok,K.C., Lilley,B.N., Kulkarni,R.V., Wingreen,N.S. and Bassler,B.L. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell*, **118**, 69–82.

59. Svenningsen,S.L., Tu,K.C. and Bassler,B.L. (2009) Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. *EMBO J.*, **28**, 429–439.

60. Castillo-Keller,M., Vuong,P. and Misra,R. (2006) Novel mechanism of *Escherichia coli* porin regulation. *J. Bacteriol.*, **188**, 576–586.

61. Joseph,B., Przybilla,K., Stuhler,C., Schauer,K., Slaghuis,J., Fuchs,T.M. and Goebel,W. (2006) Identification of *Listeria monocytogenes* genes contributing to intracellular replication by expression profiling and mutant screening. *J. Bacteriol.*, **188**, 556–568.

62. Camejo,A., Buchrieser,C., Couve,E., Carvalho,F., Reis,O., Ferreira,P., Sousa,S., Cossart,P. and Cabanes,D. (2009) *In vivo* transcriptional profiling of *Listeria monocytogenes* and mutagenesis identify new virulence factors involved in infection. *PLoS Pathog.*, **5**, e1000449.

63. Horvath,P. and Barrangou,R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science*, **327**, 167–170.

64. Georg,J., Voss,B., Scholz,I., Mitschke,J., Wilde,A. and Hess,W.R. (2009) Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol. Syst. Biol.*, **5**, 305.

65. Liu,J.M., Livny,J., Lawrence,M.S., Kimball,M.D., Waldor,M.K. and Camilli,A. (2009) Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. *Nucleic Acids Res.*, **37**, e46.

66. Romby,P. and Charpentier,E. (2010) An overview of RNAs with regulatory functions in gram-positive bacteria. *Cell Mol. Life Sci.*, **67**, 217–237.

67. Wurtzel,O., Sapra,R., Chen,F., Zhu,Y., Simmons,B.A. and Sorek,R. (2010) A single-base resolution map of an archaeal transcriptome. *Genome Res.*, **20**, 133–141.

68. Brantl,S. (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr. Opin. Microbiol.*, **10**, 102–109.

69. Wagner,E.G., Altuvia,S. and Romby,P. (2002) Antisense RNAs in bacteria and their genetic elements. *Adv. Genet.*, **46**, 361–398.
70. Weaver,K.E. (2007) Emerging plasmid-encoded antisense RNA regulated systems. *Curr. Opin. Microbiol.*, **10**, 110–116.
71. Selinger,D.W., Cheung,K.J., Mei,R., Johansson,E.M., Richmond,C.S., Blattner,F.R., Lockhart,D.J. and Church,G.M. (2000) RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nat. Biotechnol.*, **18**, 1262–1268.
72. Phan,T.T. and Schumann,W. (2009) Transcriptional analysis of the lysine-responsive and riboswitch-regulated *lysC* gene of *Bacillus subtilis*. *Curr. Microbiol.*, **59**, 463–468.
73. Loh,E., Dussurget,O., Gripenland,J., Vaitkevicius,K., Tiensuu,T., Mandin,P., Repoila,F., Buchrieser,C., Cossart,P. and Johansson,J. (2009) A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell*, **139**, 770–779.