

Supporting Information for

The global RNA-RNA interactome of *Klebsiella pneumoniae* **unveils a small RNA regulator of cell division**

Eric Ruhland, Malte Siemers, Ruman Gerst, Felix Späth, Laura Nicole Vogt, Marc Thilo Figge, Kai Papenfort, Kathrin Sophie Fröhlich

corresponding author: Kathrin Sophie Fröhlich Email: kathrin.froehlich@uni-jena.de

This PDF file includes:

Figures S1 to S12 Supporting Methodology Tables S1 to S5 SI References

Annotation and classification of transcriptional start sites (TSSs) in *K. pneumoniae* **MGH 78578***.* (A) TSSs were categorized according to their genomic positions. Proximal and distal TSSs are associated with a gene if located \leq 250 bp upstream of the start codon on the same strand. Antisense TSSs are located on the opposite strand of an annotated gene, internal TSSs within an annotated gene on the sense strand. A TSS not assigned to any of these categories is classified as orphan. (B) Distribution and overlap of classification for 3,748 TSSs identified in this study. (C) Distribution of the lengths of 5' UTRs based on proximal and distal TSSs. (D) Overlap (purple) among TSSs detected in MEP (red) and ESP (blue) distributed between the core genome (9648.1) and the five plasmids (9649.1; 9650.1; 9651.1; 9652.1; 9653.1) in *K. pneumoniae* MGH 78578*.* (E) Overlap (purple) among TSSs detected in *K. pneumoniae* MGH 78578 in this study for cells grown in LB (red) or in (1) for cells cultivated in minimal M9 medium supplemented with glucose (blue).

RNA co-immunoprecipitation and expression profiles of *K. pneumoniae***.** (A) Strategy to identify Hfq-associated RNA. RNA was co-immunoprecipitated from *K. pneumoniae* (wild-type or chromosomal *hfq::3xFLAG*) using an anti-FLAG antibody. RNA was purified, converted to cDNA, and subjected to high-throughput sequencing. (B) Growth curve of different *K. pneumoniae* strains (wild-type with [WT] or without [WT*] plasmids pKpn4 and pKpn6; chromosomal mutant *hfq::3xFLAG*; chromosomal mutant Δ*hfq*) grown in LB medium. (C) Expression analysis of Hfq-3xFLAG and conserved Hfq-dependent sRNAs. WT and mutant cells (*hfq::3xFLAG* and Δ*hfq*) were grown in LB. Total protein and total RNA samples were collected at indicated optical densities (0.25 to 2 and 3h after cells reached OD $_{600}$ of 2) and subjected to analysis via Western blotting (WB; upper panel) and Northern blotting (NB; lower panel), respectively. RNAP and 5S rRNA served as loading controls.

RIP-seq analysis of Hfq-bound sRNAs and expression profiles of selected sRNA candidates. (A) Relative distribution of different sRNA types annotated in *K. pneumoniae* (left), and distribution of different sRNA types enriched ≥3-fold in the co-IP sample compared to the WT control (right). sRNAs are categorized by their genomic location (IGR: free-standing sRNA gene in intergenic region; 5' UTR: processed from mRNA 5' end; 3' UTR: processed from mRNA 3' end; antisense: antisense to an annotated gene; IGR/as: free-standing sRNA gene in intergenic region extending into an annotated gene; 3' UTR op: free-standing sRNA gene overlapping the 3' end of an annotated gene). (B) Total RNA samples of WT and Δ*hfq* cells were collected at indicated time-points (OD_{600} of 0.25 to 2 and 3h after cells reached OD_{600} of 2) during growth in LB and subjected to Northern blot analysis. KpnR_128 is expressed from its own promoter internal to another transcriptional unit positioned upstream. Both transcripts share the same termination site and the \sim 300 nt transcript (#) is thus detected by the Kpn R128 probe. 5S rRNA served as loading control.

Relative read distribution in the RIL-seq experiment. (A) Distribution of filtered single and chimeric reads in samples recovered from the RIL-seq experiment in *K. pneumoniae* wild-type (WT) or *hfq::3xFLAG* (Hfq-3xF) cells. (B) Distribution of different RNA classes in single and chimeric reads in samples recovered from the RIL-seq experiment in *K. pneumoniae* wild-type or *hfq::3xFLAG* cells.

A

+1 dinl mRNA -35 -10 start CCTGTTGTCAACGCTAACCGGATACTGTATAAATAACCAGTATCGAACGCAGGGGGCAATTATGCGCATTGAAGTGTCTATCGCCAAAACGACCGCCTTACCTAACGGCGCCCTGGAAGCGCTTAACGCCCTGAATAAC kpn CCTGTTGTCAACGCTAACCGGATACTGTATAAATAACCAGCATCGAACGCAGGGGGCAATTATGCGCATTGAAGTGTCTATCGCCAAAACGACCGCCTTACCTAACGGCGCCCTGGAAGCGCTTAACGCCCTGAATAAC kav khs CCTGTTGTCAACGCCAACCGGATACTGTATAAATAACCAGTATCAAACGCAGGGGGCAATTATGCGCATTGAAGTGTCTATCGCCAAAACGACCGCCTTACCTAACGGCGCCTGGAAGCGCTTAACGCCCTGAATAAC kqp kaf kva CCTGTTGTCAACGCCCACCGGATACTGTATAAATAACCAGTATCAAACACAGGGGGCAATTATGCGCATCGAAGTGTCTATCGCCAAAACGACCGCCTTACCCAACGGCGCGCTGGAAGCGCTTAACC eco $sf1$ stm sty CCTGTTGTAAGCTGGCTTAATTAACTGTATAAATAACCAGTATATCAACAGGGGGCGATTATGCGTATTGAAGTCACCATTGCTAAAACTTCACCTTTGCCTGCTGCGCGCGATTGACGCGCTGGCGCGAA cro ent LexA hox DinR proc. \triangleright kpn GAACTCTCTCGCCGTATCGCCGAGCAATTCCCGGCCATCGATAGCCAGGTGAGCGTCCGCTACGCCAGCGGCAATCAGCTCTCGGTGTTTGGCGCCCTGAAAGAAGAAGAACAAAGCCGGATAAGCGAAATCC GAACTCTCTCGCCGTATCGCCGAGCAATTCCCGGCCATCGATAGCCAGGTGAGCGTCCGCTACGCCAGCGGCAATCAGCTCTCGGTGTTTGGCGCCCTGAAAGAAGACAAAGCCGGATAAGCGAAATCC kav khs GAACTCTCTCGCCGTATCGCCGAGCAATTCCCGGCCATCGATAGCCAGGTGAGCGTCCGCTACGCCACCGGTAATCAGCTCTCGGTGTTTGGCGCCCTGAAAGAAGACAAAGACCGGATAAGCGAAATCC GAACTCTCTCGCCGTATCGCCGAGCAATTCCCGGCCATCGACAGCCAGGTGAGCGTCCGCTACGCCACCGGCAATCAGCTCTCGGTGTTTGGCGCCCTGAAAGAAGACAAAGACCGGATAAGCGAAATCC kqp kaf GAACTCTCTCGCCGTATCGCCGAGCAATTCCCGGCCATCGATAGCCAGGTGAGCGTCCGCTACGCGAAACCAGCTCTCGGTGTTTGGCGCCCTGAAAGAAGACAAAGACCGGATAAGCGAAATCC kva GAACTCTCCCGGCGTATCGCCGAGCAATTCCCGGCCATCGACAGTCAGGTGAGCGTGCGCTACGCCACCGGCAACCAGCTCTCGGTGTTCGGCGCCCTGAAAGAAGACAAAGCCGGATAAGCGAAATCC GAACTTTCCCGCCGTATTCAGTATGCGTTTCCTGATAATGAAGGCCACGTATCGGTACGTTATGCCGCAGCGAATAATTTATCGGTTATTGGCGCGACAAAAGAAGATAAACAGCGCATTAGTGAAATTC ACO sfl GAACTTTCCCGCCGTATTCAGTATGCGTTTCCTGATAATGAAGGCCACGTATCGGTACGTTATGCCGCAGCGAATAATTTATCGGTTATTGGCGCAACAAAAGAAGATAAACAGCGCATTAGCGAAATTC GAACTCTCCCGCCGTATTAGCCATCATTTTCCGGAGAATTTGGGTAACGTCACCGTGCGTTACGCTACCGCCAACTATTGTCCGTCATTGGCGCATCAAAAGAGGACAAAGGAACGCATTAGCGAGATTC stm GAACTCTCCCGCCGTATTAGCCATCATTTTCCTGAGAATTTGGGTAACGTCACCGTGCGTTACGCTACCGCAACAACTTGTCCGTCATTGGCGCGTCAAAAGAGGGATAAAGAACGCATTAGCGAGATTC sty GAACTCTCTCGTCGTATTACCCACCACTTCCCCGATAACGACGGTAACGTCACCGTTATGCCGCTGCGAATAATTTGTCCGTCATCGGCGCGACAAAAGAAGATAAACAGCGTATCAGCGAAATTC cro GAATTATCCCGTCGTATCAATGACAATTTCCCCGACCAGAGCGGAGCCGTAACCGCTACGCCACAGCCAATAATCTCCCCCATCATGGGTGCAAAAGACGATAAAGAACGCATTACCGAGATTC ent

Conservation of *dinI* **in enterobacteria.** (A) Multiple alignment of the *dinI* gene of different enterobacterial species (kpn: *Klebsiella pneumoniae* MGH 78578; kqv: *Klebsiella quasivariicola* 08A119; khs: *Klebsiella pneumoniae subsp. pneumoniae* HS11286; kqp: *Klebsiella quasipneumoniae* KqPF26; kaf: *Klebsiella africana* 200023; kva: *Klebsiella variicola* LEMB11; eco: *Escherichia coli* MG1655; sfl: *Shigella flexneri* 301; stm: *Salmonella* Typhimurium LT2; sty: *Salmonella typhi* CT18; cro: *Citrobacter rodentium* ICC168; ent: *Enterobacter sp.* 638) was calculated using the MultAlin tool (2). Nucleotides are colored regarding their degree of conservation (red: high conservation; blue: partial conservation; black: little or no conservation). The -10 and -35 elements of the *dinI* promoter, as well as the start and stop codons and the Rho-independent terminator are boxed, the transcriptional start site (as determined for *K. pneumoniae*; compare Fig. S6) is indicated by an arrow. The 5' ends of the two most abundant processing intermediates (proc. and DinR) from *dinI* mRNA in *K. pneumoniae* MGH 78578 are indicated by open triangles. The LexA box of *dinI* (as determined for *E. coli* in (3, 4)) is highlighted in orange. (B) Multiple alignment of the DinI protein of different enterobacterial species (as in (A)) was calculated using the MAFFT tool (5).

Determination of the 5' ends of *dinI* **mRNA and its major processing products.** (A) Readmappings of dRNA-seq (-TEX: black traces; +TEX: red traces) and Hfq RIP-seq (control: light blue traces; Hfq-3xFLAG: dark blue traces) to the *dinI* locus. The y-axes in individual experiments were set to the same scale. The positions on the *K. pneumoniae MGH 78578* genome are indicated above the sequencing traces, and the *dinI* CDS is indicated by a grey arrow. The TSS of *dinI* was determined by analysis of the dRNA-seq dataset, and is marked by a black arrow. Stable processing products were identified from dRNA-seq and Hfq RIP-seq and sizes are below the panels. (B) Primer extension analysis of *dinI* mRNA (using a genespecific primer). RNA was extracted from wild-type cells at $OD₆₀₀$ of 2.0, and 30 min after DNA damage was induced through mitomycin C (MMC) or ultraviolet light (UV). RNA purified from *dinIR* mutant cells served as a control. The 5' ends of *dinI* mRNA (black arrowhead) and its processing products (proc. and DinR; open arrowheads; DinR-S is too close to the 3' end to be resolved in this experiment) were mapped using a sequencing ladder, and correspond to the sites determined by dRNA-seq. (C) Sequence at the *dinI* locus of *K. pneumoniae MGH 78578*. The -35 and -10 elements of the *dinI* promoter are boxed, the LexA box is high-lighted in orange. The 5' ends of *dinI* mRNA (black arrowhead) and its processing products (proc., DinR, DinR-S; open arrowheads) are indicated. Numbering refers to the TSS as position +1.

Expression of *dinI* **mRNA and DinR sRNA in** *K. pneumoniae***.** (A) RNA samples were collected at different time-points over growth (OD₆₀₀ from 0.25 to 2.0), and 30 min after cells had reached an OD₆₀₀ of 2.0 in the absence (*ctrl.*) or presence of DNA damage induced through mitomycin C (MMC), ciprofloxacin (CPX) or ultraviolet light (UV). Expression of *dinI* mRNA and DinR was assessed by Northern blot analysis; 5S rRNA served as loading control. (B) Quantification of *dinI* mRNA and DinR sRNA signals relative to the expression at OD₆₀₀ of 2.0.

DinR is processed from the *dinI* **mRNA.** (A) *K. pneumoniae* WT or ∆*dinIR* cells carrying either an empty control vector (pBAD_{KP}-*ctrl.*) or plasmids harbouring different variants of the *dinI* gene (either starting from the TSS [pTSS_*dinIR*] or 50 bp upstream of the TSS [pP_{own} *dinIR*]) were grown to $OD₆₀₀$ of 2.0, when DNA damage was induced by the addition of MMC. DinR and *dinI* mRNA levels were determined by Northern blot analysis of RNA samples collected at the indicated time-points. 5S ribosomal RNA served as a loading control. (B) *Salmonella* Typhimurium expressing a temperature-sensitive RNase E variant *(rne-*TS) or an isogenic control strain (*rne-ctrl.*) were transformed with pBAD_{EC}-dinIR and grown to OD₆₀₀ of 0.25 at the non-permissive temperature of 30 °C. Cultures were split and cultivation was continued for 45 min at 30 °C or at 44 °C to inactivate RNase E. RNA samples were collected prior to (-) and 15 min after (+) induction of *dinIR* expression with arabinose. Expression of *dinI* mRNA and DinR were determined by Northern blot analysis. 5S ribosomal RNA served as a loading control, and accumulating precursors of 5S rRNA are indicated by asterisks.

Regulation of *ftsZ* **expression and conservation of the 5' UTR.** (A) Quantification of FtsZ protein levels. FtsZ expression was determined by Western blot analysis using a FtsZ-specific antiserum in total protein samples collected from *K. pneumoniae*, *E. coli* and *V. cholerae* carrying either an empty control vector (pBAD_{KP}-ctrl., pBAD_{EC}-ctrl., or pBAD_{VC}-ctrl., respectively) or pBAD variants for the expression of $DimR$ (pBAD_{KP}-DinR, pBAD_{EC}-DinR, or pBADVC-DinR, respectively). Bacteria were diluted from overnight cultures into fresh medium, and sRNA expression was induced by the addition of arabinose for 5 hours. FtsZ levels in the control sample (- L-Ara) was set to 1; error bars denote the standard deviation from five biological replicates. (B) Non-redundant alignment of the *ftsZ* 5' UTR of different bacterial species (kpn: *Klebsiella pneumoniae* MGH 78578; cko: *Citrobacter koseri* ATCC BAA-895; eco: *Escherichia coli* MG1655; sfl: *Shigella flexneri* 301; stm: *Salmonella* Typhimurium LT2; vch: *Vibrio cholerae* O1 El Tor N16961) was calculated using the Multalign tool (2). The stop codons of the upstream gene *ftsA* and the start codons of *ftsZ* are underlined, two transcriptional start sites (as determined by dRNA-seq in *K. pneumoniae*) are indicated by arrows. All nucleotides are colored with regard to their degree of conservation (red: high conservation; blue: partial conservation; black: little or no conservation).

In vitro **analysis of DinR base-pairing on** *ftsZ* **mRNA.** Structure probing of 5′ end-labelled *ftsZ* mRNA (TSS at –58 to +60 relative to the start codon; 0.4 pmol) with RNase T1 (lanes 4 to 11) and lead(II) acetate (lanes 12 to 19) in the absence or presence of 2 pmol (5x) Hfq protein and 4 pmol (10x) sRNAs DinR (WT), DinR-S (S), or DinR variants (M1-M4; compare 5(B)). RNase T1 and alkaline ladders of the *ftsZ* transcript were used to map the positions of individual nucleotides. The positions of G residues are indicated relative to the translational start site. Putative DinR binding sites are marked in blue.

DinR overexpression causes cell filamentation. (A) *K. pneumoniae* carrying either an empty control vector (pBAD_{KP}-*ctrl.*) or pBAD_{KP} variants for the expression of DinR, DinR-S, DinR-M1, DinR-M2, DinR-M3 or DinR-M4 were diluted from overnight cultures into fresh medium, and sRNA expression was induced by the addition of arabinose. Cell morphology was assessed by phase contrast microscopy after 5 hours. (B) Analysis of cell lengths in samples described in (A). The center line indicates the median, boxes represent the 25th and 75th percentiles, and lower and upper whiskers represent the 10th and 90th percentiles, respectively. (C) *E. coli* carrying the post-transcriptional *ftsZ::gfp* reporter and either pBAD_{EC}-ctrl. or pBAD_{EC} variants for the expression of DinR, DinR-S, DinR-M1, DinR-M2, DinR-M3 or DinR-M4 were grown to OD₆₀₀ of 0.5. Total RNA samples were prepared from cells collected prior to (-) or 10 min after the addition of arabinose (+). DinR and *ftsZ::gfp* mRNA levels were determined by Northern blot analysis; 5S ribosomal RNA served as a loading control.

Complementation of the *dinIR* **mutant.** (A) Schematic of the *dinIR* locus of *K. pneumoniae* in wild-type and mutant cells. In the chromosomal mutant Δ*dinIR* deletion, a 81 bp scar replaces the sequence from the TSS to the end of the 3' UTR. In the chromosomal mutant Δ*dinR*, a 81 bp scar replaces the sequence from the stop codon to the beginning of the terminator hairpin (nt 84 to 126 of the sRNA). The size of the *dinI* mRNA increases to 386 nt. (B) Schematic of complementation constructs of *dinIR*. Plasmid pPown_*dinIR* expresses *dinIR* under control of its own promoter; DinI protein is produced. Plasmid pPown_*dinI*AGG*R* expresses *dinIR* under control of its own promoter; the stop codon of *dinI* is changed from ATG to AGG; DinI protein is not produced. Plasmid pP_{own} *dinIR*M16 expresses *dinIR* under control of its own promoter; nt 114 to 130 of DinR have been deleted; DinI protein is produced. (C) *K. pneumoniae* wild-type cells or *dinIR* mutant cells carrying complementation plasmids (as described in (B)) were diluted from overnight cultures into fresh medium, split and either treated with MMC (1 µg/mL) to induce DNA damage or left untreated. Cell morphology was assessed by phase contrast microscopy after 5 hours and quantified. The center line indicates the median, boxes represent the 25th and 75th percentiles, and lower and upper whiskers represent the 10th and 90th percentiles, respectively. (D) RNA samples prepared from cells treated as described in (C) were analyzed on Northern blot; 5S ribosomal RNA served as a loading control.

SUPPORTING METHODOLOGY

Construction of Bacterial Strains

Klebsiella pneumoniae subsp. pneumoniae MGH 78578, *E. coli* Top10 and *V. cholerae* C6706 are referred to as wild-type strains in this study.

Plasmids were introduced in *K. pneumoniae* and *E. coli* by electrotransformation, and into *V. cholerae* by conjugation via the *E. coli* S17 helper strain using standard protocols. To construct chromosomal mutations, we followed a published procedure based on the widely-used *E. coli* one-step inactivation method based on the λ-Red recombinase system (6). Briefly, wild-type cells carrying pACBSR-Hyg (encoding the λ-Red genes under control of the *araBAD* promoter) were transformed with PCR products (obtained by overlap extension PCR) containing the *aac(3)IV* gene (conferring apramycin resistance) flanked by FRT recombination sites and ~500 bp of sequences homologous to the genomic integration sites. For fragments carrying the *aac(3)IV* marker gene, pIJ773 (KFO-1407/KFO-1376) or pER15 (KFO-1375/KFO-1376) were amplified and combined with individual flanking regions (∆*hfq*: KFO-1252/KFO-1408 and KFO-1255/KFO-1378; *hfq*-3xFLAG: KFO-1252/KFO-1377 and KFO-1255/KFO-1378; ∆*sulA*: KFO-1798/KFO-1800 and KFO-1799/KFO-1801; ∆*dinI*: KFO-1792/KFO-1821 and KFO-1905/KFO-1793; ∆*dinR*: KFO-1792/KFO-2001 and KFO-2000/KFO-1793). Transformants were recovered in the presence of apramycin, and desired mutations were confirmed by colony PCR and sequencing. The temperature-sensitive pACBSR-hyg was lost during growth in LB at 37 °C. To eliminate the *aac(3)IV* cassette, the resulting strain was transformed with pFLP-hyg, and expression of the FLP recombinase was induced with arabinose. Recombination at the FRT site and correct excision of the cassette was confirmed by colony PCR and sequencing, and the helper plasmid was removed by passaging cells at 37 °C. Multiple mutations were introduced by sequentially performing the genomic modifications described above.

Unexpectedly, introduction of the pACBSR-Hyg helper plasmid resulted in selective loss of accessory plasmids pKPN4 and pKPN6, respectively, from a subset of transformed wild-type *K. pneumoniae* MGH 78578, a phenotype that had not been reported in the original protocol. We thus verified that neither loss of the two plasmids nor addition of the epitope to the RNAbinding protein expressed from its native promoter in MGH 78578 impaired bacterial growth under standard conditions when compared to WT cells (Fig S2B). In addition, we performed phenotypic characterization of our results throughout the manuscript using an isogenic WT strain, in which pKPN4 and pKPN6 had been likewise removed.

Bacterial growth conditions

K. pneumoniae, Escherichia coli and *V. cholerae* strains were grown aerobically at 37 °C in LB broth or on LB agar. Where appropriate, media were supplemented with kanamycin (50 µg/mL), chloramphenicol (20 µg/mL), ampicillin (100 µg/mL), hygromycin (100 µg/mL), or apramycin (50 µg/mL). Expression from the araBAD promoter was induced with L-arabinose (final concentration: 0.2%). Unless stated otherwise, DNA damage was triggered in *K. pneumoniae* by adding mitomycin C (MMC; 0.5 µg/mL) or ciprofloxacin (CPX; 4 µg/mL) to the culture, or by irradiation with UV light ($λ=254$ nm; 10 mJ/cm²).

Plasmid construction

All plasmids and oligonucleotides used in this study are listed in Tables S4 and S5.

A template plasmid (pER15) for the chromosomal integration of the 3xFLAG epitope tag was constructed via PCR amplification of pIJ773 (KFO-1250/KFO-1251), and self-ligation was carried out as in (7).

The translational GFP reporter fusion of *ftsZ* (pFS7) was constructed via Gibson assembly (GA; NEB #E2611L) as recommended by the manufacturer using a PCR product amplified from *K. pneumoniae* genomic DNA (KFO-1713/KFO-1714) and the linearized pXG10 vector (KPO-1702/KPO-1703). For the construction of the *ftsZ(M3)::gfp* variant (pER44), plasmid pFS07 was used as a template in PCR amplification with primer set KFO-1851/KFO-1852, and the obtained fragment was self-ligated.

For plasmids expressing DinR from the arabinose-inducible pBAD plasmid, inserts amplified from *K. pneumoniae* genomic DNA were cloned into linearized vectors via GA. For expression in *E. coli*, backbone pKP8-35 was amplified via KPO-0196/KPO-0411 and combined with an insert amplified with KFO-1718/KFO-1716. For expression in *V. cholerae*, backbone pBAD1K was amplified via KPO-0196/KPO-1397 and combined with an insert amplified with KFO-1715/KFO-2140. For expression in *K. pneumoniae*, an apramycin resistance cassette (amplified from pIJ773 with KFO-1830/KFO-1831) was introduced into pKP8-35 (linearized with KFO-1832/KFO-1833), and the obtained plasmid (pER41) was amplified via KPO-0196/KPO-0411 and combined with an insert amplified with KFO-1718/KFO-1716 by GA.

For plasmids expressing variants of DinR, pFS1 or pER41 served as a template for PCR amplification with primer pairs KFO-1992/KFO-1993 (pER66 and pER90); KFO-1892/KFO-1893 (pFS51 and pER85); KFO-1849/KFO-1850 (pER43 and pER84); KFO-1990/KFO-1991 (pER65 and pER89), and obtained fragments were self-ligated.

To verify transcriptional control of DinR expression, two variants of *dinI* lacking (KFO-2028/KFO-1716; pER82) or including the annotated promoter (KFO-2029/KFO-1716; pER83) were inserted via GA into linearized pER41 lacking the *araBAD* promoter (amplified by KFO-2027/KPO-0411). Plasmid pER83 served as template for PCR amplification with primer pairs KFO-2719/KFO-2720 to obtain pER156 and with KFO-2721/KFO-2723 to obtain pER158.

T7 transcription and 5' end-labelling of RNA

RNA was *in vitro* synthesized and 5' end-labelled as described before (8, 9). In short, DNA templates carrying the T7 promotor were amplified by PCR using the oligonucleotides listed in Table S5. *In vitro* transcription of RNA from template DNA was performed using the AmpliScribe T7-Flash transcription kit (Epicentre). RNA was dephosphorylated using calfintestinal alkaline phosphatase (NEB), subsequently extracted with P:C:I (phenol/chloroform/isoamyl alcohol; 25:24:1) and ethanol precipitated. 5' end labelling was achieved by incubation of dephosphorylated RNA with [32P]- γATP and polynucleotide kinase (NEB) for 1 h at 37 °C. Unincorporated nucleotides were removed using Amersham MicroSpin G-50 columns (Cytiva). Labelled RNA was then purified on a denaturing 6 % PAA / 7 M Urea gel, eluted with RNA elution buffer (0.1 M sodium acetate, 0.1 % SDS, 10 mM EDTA) at 4 °C overnight and recovered by P:C:I extraction.

Riboprobes

The riboprobes to detect DinR, *dinI* and *sulA* mRNAs were synthesized by T7-mediated *in vitro* transcription of ∼200 ng of template DNA (amplified on *Klebsiella* gDNA with KFO-1762/KFO-1763; KFO-2554/KFO/2555; KFO-2288/KFO-2260) in the presence of [32P]-α-UTP with the MAXIscript T7 transcription kit (Invitrogen). Unincorporated nucleotides were removed using MicroSpin G-50 columns (Cytiva).

Primer extension analysis

For primer extension, 5 μg of RNA were denaturated in the presence of 1 pmol 5′ end-labelled primer (KFO-1745) at 70°C for 2 min and adjacently chilled on ice for 5 min. Next, the samples were mixed with the reaction mix (1X first strand buffer, 5 mM DTT, 0.5 mM each dATP, dGTP, dCTP and dTTP) at 42°C, and SuperScript III (100 U; Invitrogen) was added. cDNA synthesis was performed at 50°C for 60 min, followed by incubation at 70°C for 15 min to inactivate the enzyme. Samples were treated with RNase H (2.5 U) for 15 min at 37°C and the reaction was stopped by the addition of GLII loading buffer. Samples was separated by electrophoresis on a 6% sequencing gel together with a template-specific ladder (prepared using the SequiTherm EXCELII DNA Sequencing Kit).

dRNA sequencing

In brief, RNA samples were fragmented using ultrasound, followed by treatment with polynucleotide kinase (PNK). For the depletion of processed transcripts and the control reaction, equal amounts of RNA were incubated with or without terminator 5′-phosphatedependent exonuclease (TEX; Lucigen). RNA samples were poly(A)-tailed using poly(A) polymerase and 5′-triphosphates were removed by applying tobacco acid pyrophosphatase (TAP). An RNA adapter was ligated to the 5' end of the RNA prior to cDNA synthesis using an oligo(dT) primer and M-MLV-RNase H- reverse transcriptase. Upon amplification of the cDNA with a high-fidelity polymerase and gel fractionation, the resulting libraries were pooled and sequenced on an Illumina NextSeq 500 system with 1x75 bp read length.

Sequencing reads from the dRNA-seq experiment were processed and analysed according to (10) with minor modifications. Briefly, for every replicate in both conditions, the coverage and based on it, the difference between neighboring nucleotides was computed. Local maxima in the differences were called as TSS if the difference between maximum and minimum coverage in a 7 nt window around the TSS was above 3 TPM, and the ratio between the maximum and the minimum was above 1.3. The set of putative TSS was manually curated and the resulting predictions were used to annotate genes with a 5' UTR within a window of 250 nt upstream of an annotated gene. The most distant TSS was then used to define the 5' UTR of the corresponding gene. 3' UTRs were annotated by extending an area downstream of a gene, up to 150 nt in length unless the next annotation was closer than that window. Custom scripts are available from GitHub (https://github.com/maltesie/KpnInteractomePaper; DOI: 10.5281/zenodo.8409716).

Hfq RIP-seq analysis

Demultiplexed raw reads were imported into the CLC Genomics Workbench (Qiagen) and subjected to quality control and adaptor trimming. The trimmed reads were mapped to the *Klebsiella pneumoniae subsp. pneumoniae* MGH 78578 reference genome (NCBI accession numbers NC_009648.1, NC_009649.1, NC_009651.1, and NC_009653.1) with standard parameter settings. Annotations of sRNAs were added manually. Fold enrichment in samples expressing Hfq-3xFLAG over the untagged control samples was calculated using the CLC "Differential Expression for RNA-Seq" tool.

RIL-seq analysis

Briefly, cells corresponding to 40 OD₆₀₀ units were subjected to UV crosslinking of proteins and RNA followed by cell lysis and co-immunoprecipitation using a monoclonal anti-FLAG antibody (Sigma; #F1804). Recovered RNA was trimmed by RNase A/T1 treatment, and proximal RNAs were ligated. Upon proteinase K treatment, RNA was extracted, fragmented and DNase digested. A previously published protocol (11) was adapted for the strain-specific depletion of ribosomal RNA. Co-immunoprecipitated RNA samples were mixed with an oligonucleotide mix (final concentrations: 5.8 nM for 16S and 23S oligos; 11.6 nM for 5S oligos) in the presence of 1x SSC and 1 mM EDTA. The mixture was denatured, cooled down and incubated with streptavidin beads (ThermoFisher; #65001) in 0.5x SSC at room temperature and 50°C for 5 min each, respectively. Depleted RNA in the supernatant was purified using AMPure XP beads (Beckman-Coulter; #A63881) and subjected to cDNA library preparation. Upon PCR amplification, libraries were sequenced in paired-end mode on an Illumina NextSeq 1000 with 150 bp read-length.

The sequence file was demultiplexed using the split libs command from RNASeqTools (DOI: 10.5281/zenodo.8388882) and the resulting files were analyzed with ChimericFragments (DOI: 10.5281/zenodo.8376810) used with the default parameters set except for the following: min seed length=15, autocomplete utrs=false, bp shift weight=0.0, max bp fdr=0.35 and min reads=5 as described in (12).

Annotation of sRNAs

Complementation of sRNA annotation was based on several criteria. Conserved enterobacterial sRNAs were identified by comparison with published datasets (13-16). Candidate sRNAs in IGRs were selected if a clear TSS, a length ranging from 50 to 350 nts, and a termination site were identified; sequences were scanned for putative ORFs $(≥ 20$ aa). 5' UTR-derived sRNAs were annotated based on the detection of a processing site or a sharp drop in coverage upstream of the translational start site of the associated mRNA. 3' UTRderived sRNAs were selected based on the presence of a processing site or a TSS within the 3' end of the associated mRNA and a shared termination site.

Identification of LexA motifs

To identify previously unknown sRNAs contributing to the regulation of the SOS response we examined the Hfq RIP-seq dataset for potentially LexA-controlled transcripts, exploiting our accurate annotation of TSSs by dRNA-seq. We used MEME (17) to compute a LexA position weight matrix based on the sequences of 21 known SOS boxes of *E. coli* (3), and then searched for the obtained motif (WACTGTATATWHAHMCAGTA) within *K. pneumoniae* transcription initiation sites (-70 to +30 relative to the TSS). We identified putative LexA binding sites preceding 71 transcription units (see Table S2).

Chromatin immunoprecipitation (ChIP)

ChIP was performed following the previously published procedures (4, 10) with minor modifications. In brief, *K. pneumoniae* MGH 78578 wild-type cells were grown in two biological replicates in LB medium to a final $OD₆₀₀$ of 2.0. The culture was split, and incubation was continued for an additional 30 min in the presence or absence of MMC (1 µg/mL).

Formaldehyde was added at a final concentration of 1% to cross-link DNA and proteins, and the reaction was quenched after 20' by the addition of glycine (0.5 M). Cells were washed in 1x TBS, and lysed in lysis buffer (50 mM Hepes-KOH pH=7.0, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS) containing 4 mg/ml lysozyme. DNA was sheared to an average size of ~200 bp using a BioRuptor (5x 30 s pulses). Insoluble material was removed by centrifugation, and a sample of the supernatant was stored as input control. The samples were mixed with protein A magnetic beads (Thermo Fisher #10001D), split and either incubated with anti-LexA antibody (Sigma #06-719; +AB) or no antibody (-AB) at 4°C overnight. Samples were washed twice in lysis buffer, once in modified lysis buffer (500 mM NaCl), once in ChIP wash buffer (10 mM Tris-HCl pH=8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet-P40, 0.5% sodium deoxycholate), and once in TE. Samples were eluted in elution buffer (50 mM Tris-HCl, 1 mM EDTA, 1% SDS) at 65°C, and cross-linking was reverted by boiling samples (including the input) for 10 min. DNA was purified by P:C:I extraction, and analyzed by quantitative PCR (qPCR) using the GoTaq qPCR Master Mix (Promega #A6002) on a CFX96 Real-Time PCR system (Bio-Rad), with *sgrR* CDS as a control.

Image analysis

The pipeline detects the bacteria and includes a background subtraction, noise removal and automated thresholding. To prevent clustering of cells in close vicinity, a second branch segments the halo around each object and subtracts those from the results of the cell thresholding. To resolve overlaps and measure cell lengths, the segmented objects are transformed into a graph structure that contains information about the cell's medial axis and diameter. This representation allowed the usage of an algorithm to split overlapping clusters into individual objects by the removal of junction vertices and addition of alternative edges minimize the cell curvature. The length of each graph component and thus the bacterial cell is determined by a JIPipe-provided operation.

A detailed explanation of the workflow and its individual components, all results as well as the utilized software versions, is available through https://asbdata.hkijena.de/RuhlandEtAl2023_PNAS.

Supplementary Table S1 – sRNAs in *K. pneumoniae* **MGH 78578**

(a) determined for newly annotated sRNA candidates

Supplementary Table S2 – LexA boxes in MGH 78578

Supplementary Table S3 – Bacterial strains

Supplementary Table S4 – Plasmids

Supplementary Table S5 – Oligonucleotides

Supplementary References

- 1. Kim D*, et al.* (2012) Comparative analysis of regulatory elements between Escherichia coli and Klebsiella pneumoniae by genome-wide transcription start site profiling. *PLoS Genet* 8(8):e1002867.
- 2. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16(22):10881-10890.
- 3. Fernandez De Henestrosa AR*, et al.* (2000) Identification of additional genes belonging to the LexA regulon in Escherichia coli. *Mol Microbiol* 35(6):1560-1572.
- 4. Wade JT, Reppas NB, Church GM, & Struhl K (2005) Genomic analysis of LexA binding reveals the permissive nature of the Escherichia coli genome and identifies unconventional target sites. *Genes Dev* 19(21):2619-2630.
- 5. Katoh K, Rozewicki J, & Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20(4):1160-1166.
- 6. Huang TW*, et al.* (2014) Capsule deletion via a lambda-Red knockout system perturbs biofilm formation and fimbriae expression in Klebsiella pneumoniae MGH 78578. *BMC Res Notes* 7:13.
- 7. Sharma CM, Darfeuille F, Plantinga TH, & Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* 21(21):2804-2817.
- 8. Fröhlich KS, Papenfort K, Fekete A, & Vogel J (2013) A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J* 32(22):2963-2979.
- 9. Papenfort K*, et al.* (2006) sE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay. *Mol Microbiol* 62(6):1674- 1688.
- 10. Huber M*, et al.* (2022) An RNA sponge controls quorum sensing dynamics and biofilm formation in Vibrio cholerae. *Nat Commun* 13(1):7585.
- 11. Culviner PH, Guegler CK, & Laub MT (2020) A Simple, Cost-Effective, and Robust Method for rRNA Depletion in RNA-Sequencing Studies. *mBio* 11(2).
- 12. Siemers M, Lippegaus A, & Papenfort K (2023) ChimericFragments: Computation, analysis, and visualization of global RNA networks. *bioRxiv* DOI: 10.1101/2023.12.21.572723.
- 13. Bar A, Argaman L, Altuvia Y, & Margalit H (2021) Prediction of Novel Bacterial Small RNAs From RIL-Seq RNA-RNA Interaction Data. *Front Microbiol* 12:635070.
- 14. Bruchmann S*, et al.* (2015) Deep transcriptome profiling of clinical Klebsiella pneumoniae isolates reveals strain and sequence type-specific adaptation. *Environ Microbiol* 17(11):4690-4710.
- 15. Hör J, Matera G, Vogel J, Gottesman S, & Storz G (2020) Trans-Acting Small RNAs and Their Effects on Gene Expression in Escherichia coli and Salmonella enterica. *EcoSal Plus* 9(1).
- 16. Iosub IA*, et al.* (2021) The mRNA derived MalH sRNA contributes to alternative carbon source utilization by tuning maltoporin expression in E. coli. *RNA Biol* 18(6):914-931.
- 17. Bailey TL & Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2:28-36.
- 18. Thelin KH & Taylor RK (1996) Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by Vibrio cholerae O1 El Tor biotype and O139 strains. *Infect Immun* 64(7):2853-2856.
- 19. Figueroa-Bossi N, Valentini M, Malleret L, & Bossi L (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Gene Dev* 23(17):2004-2015.
- 20. Hoyos M, Huber M, Forstner KU, & Papenfort K (2020) Gene autoregulation by 3' UTR-derived bacterial small RNAs. *Elife* 9.