Silencing of natural transformation by an RNA chaperone and a multitarget small RNA

Laetitia Attaiech[a,b,c,d,e,f,g], Aïda Boughammoura[a,b], Céline Brochier-Armanet[h], Omran Allatif[c,d,e,f,g], Flora Peillard-Fiorente[a,b], Ross A. Edwards[i], Ayat R. Omar[i], Andrew M. MacMillan[i], Mark Glover[i], and Xavier Charpentier[a,b,c,d,e,f,g,1]

[a]CNRS UMR5240, Microbiologie Adaptation et Pathogénie, 69622 Villeurbanne, France; [b]Université Claude Bernard Lyon 1, 69100 Villeurbanne, France; [c]International Center for Infectiology Research (CIRI), International Center for Infectiology Research, Université de Lyon, 69364 Lyon, France; [d]INSERM U1111, 69364 Lyon, France; [e]Ecole Normale Supérieure de Lyon, 69364 Lyon, France; [f]Université Claude Bernard Lyon 1, Centre International de Recherche en Infectiologie, 69364 Lyon, France; [g]CNRS UMR5308, 69007 Lyon, France; [h]Laboratoire de Biométrie et Biologie Evolutive, CNRS UMR5558, Université de Lyon, Université Lyon 1, F-69622 Villeurbanne, France; and [i]Department of Biochemistry, University of Alberta, Edmonton, AB T6G 2H7, Canada

Edited by Ralph R. Isberg, Howard Hughes Medical Institute/Tufts University School of Medicine, Boston, MA, and approved May 26, 2016 (received for review January 29, 2016)

A highly conserved DNA uptake system allows many bacteria to actively import and integrate exogenous DNA. This process, called natural transformation, represents a major mechanism of horizontal gene transfer (HGT) involved in the acquisition of virulence and antibiotic resistance determinants. Despite evidence of HGT and the high level of conservation of the genes coding the DNA uptake system, most bacterial species appear non-transformable under laboratory conditions. In naturally transformable species, the DNA uptake system is only expressed when bacteria enter a physiological state called competence, which develops under specific conditions. Here, we investigated the mechanism that controls expression of the DNA uptake system in the human pathogen *Legionella pneumophila*. We found that a repressor of this system displays a conserved ProQ/FinO domain and interacts with a newly characterized *trans*-acting sRNA, RocR. Together, they target mRNAs of the genes coding the DNA uptake system to control natural transformation. This RNA-based silencing represents a previously unknown regulatory means to control this major mechanism of HGT. Importantly, these findings also show that chromosome-encoded ProQ/FinO domain-containing proteins can assist *trans*-acting sRNAs and that this class of RNA chaperones could play key roles in post-transcriptional gene regulation throughout bacterial species.

natural transformation | RNA chaperone | non-coding RNA | *Legionella pneumophila* | ProQ/FinO

Natural transformation is a common mode of horizontal gene transfer in bacteria. It results from the intrinsic capacity of bacteria to import exogenous DNA and integrate it by recombination in their chromosome (1). Active acquisition of random parts of the genetic material released by phylogenetically close organisms produces genetic polymorphism and functions as "localized sex" in reference to the function of sex in eukaryotic organisms (2). In addition, evidence suggests that genetic transformation can also occur with distantly related and even damaged DNA (3, 4). Depending on the bacteria's biotope, the constant and random acquisition of genetic material could be as harmful as it is beneficial, and in most transformable species, natural transformation is a strictly regulated process (1, 5). Natural transformation only occurs when bacteria enter a specialized physiological state, called competence, during which a DNA uptake system is set up (6). This system generally involves a type IV pilus exposed at the cell surface and a transporter associated to the cytoplasmic membrane (1, 6). The type IV pilus is thought to initially interact with DNA and convey it to the small ComEA protein that binds double-strand DNA (dsDNA) (7). The captured dsDNA is then converted into single-strand DNA (ssDNA) and transported across the cytoplasmic membrane through a transmembrane channel formed by ComEC (8). The ssDNA entering the cytoplasm is rapidly loaded with ssDNA binding proteins SsbB and DrpA and can recombine with the chromosome (9, 10).

The concerted expression of the DNA uptake system during competence invariably relies on transcriptional activation (1). In Gram-positive bacteria, it results from the action of a transcriptional activator (e.g., ComK in *Bacillus subtilis*) or sigma factors (e.g., σ^X in *Streptococcus pneumoniae*, σ^H in *Staphylococcus aureus*). In the Gram-negative *Haemophilus influenzae* and *Vibrio cholerae*, it involves the transcription activator TfoX/Sxy. However, a number of species that are naturally transformable lack these known competence activators. One such species is the Gram-negative pathogen *Legionella pneumophila*, which was found to develop competence under microaerophilic growth, exposure to DNA-damaging agents, or suboptimal growth temperature (11–13). Possible competence regulatory elements in *L. pneumophila* were revealed by a genetic screen that identified a gene annotated "*proQ*, activator of osmoprotectant ProP" as a repressor of natural transformation (14). The *proQ* gene was initially identified in *Escherichia coli* as an activator of the proline transporter ProP (15), but recent evidence suggests that this effect is indirect (16). More significantly, ProQ bears the conserved PFAM domain 04352 (named the ProQ/FinO domain), which has been functionally and structurally studied in the F plasmid-encoded FinO protein (17, 18). FinO is an RNA chaperone that down-regulates the conjugative transfer of IncF plasmids

Significance

Natural transformation is a major mechanism of horizontal gene transfer (HGT) by which bacteria take up exogenous DNA directly in their environment and integrate it in their genome. Acquiring new genetic information may confer an adaptive advantage but an uncontrolled uptake of foreign DNA may be harmful. We document a previously unsuspected means to control HGT by natural transformation in the human pathogen *Legionella pneumophila*. We found that the DNA uptake system required for natural transformation is subjected to silencing. A member of the widespread ProQ/FinO domain-containing protein family acts as an RNA chaperone and allows the targeting of the mRNAs of the genes coding the DNA uptake system by a newly identified *trans*-acting small RNA.

Author contributions: L.A., A.B., C.B.-A., A.M.M., M.G., and X.C. designed research; L.A., A.B., C.B.-A., F.P.-F., R.A.E., A.R.O., and X.C. performed research; O.A. contributed new reagents/analytic tools; O.A. performed statistical analyses; L.A., A.B., C.B.-A., R.A.E., A.R.O., A.M.M., M.G., and X.C. analyzed data; and L.A., C.B.-A., M.G., and X.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: RNA-seq data were deposited to the MINSEQE-compliant public ArrayExpress database, https://www.ebi.ac.uk/arrayexpress (accession no. E-MTAB-4094) and brokered to the European Nucleotide Archive, www.ebi.ac.uk (accession no. ERP013398).

[1]To whom correspondence should be addressed. Email: xavier.charpentier@univ-lyon1.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601626113/-/DCSupplemental.

between *Enterobacteriaceae* by controlling the function of an antisense RNA, FinP (19). FinO facilitates the interaction between FinP and the complementary 5′ untranslated region (UTR) of the mRNA of *traJ* (20, 21). Sense–antisense pairing prevents translation of the TraJ transcriptional activator of the plasmid *tra* operon and thereby inhibits plasmid conjugation. Although *finO* is restricted to IncF plasmids, genome-encoded proteins with a ProQ/FinO domain have been identified, but their function remains elusive (19). As the *L. pneumophila* ProQ represses natural transformability (14) and to avoid any unfounded inference with proline metabolism, we will refer to it by its locus tag name in the Paris strain, Lpp0148.

We characterize here Lpp0148 and describe its biological function in interaction with the first *trans*-acting sRNA-negative regulator of competence. We document an entirely novel mechanism of competence regulation by post-transcriptional silencing, diverging from the current notion that competence strictly depends on transcriptional activation. Our study expands the gene regulatory functions of ProQ/FinO domain-containing proteins by demonstrating that this class of RNA chaperones can act on regulatory *trans*-acting sRNAs.

Results

The ProQ/FinO Domain-Containing Protein Lpp0148 Specifically Controls a Competence Regulon. To determine if Lpp0148 is a specific repressor of competence, we analyzed the global transcriptional activity in a *lpp0148* mutant created by the introduction of a premature stop codon (denoted $lpp0148_{TAA}$) (22). The $lpp0148_{TAA}$ mutant, which is highly transformable (Fig. S1), was grown to exponential growth phase (OD_{600} of 0.8) and subjected to RNA-seq transcriptional profiling (Table S1). The regulon controlled by Lpp0148 (fold change > 2, $P < 0.01$) consists of 11 genes upregulated in the mutant strain and arranged in seven potential transcriptional units (Table S1). Among those genes, six are homologous to either genes encoding elements of the DNA uptake system (*comEA*, *comEC*, and *comF*) or genes previously found induced in competent bacteria (*comM*, *radC,* and *mreB*) (23–25). *lpp1976* is the last gene of an operon with *lpp1977* and *lpp1978*, both of which are moderately induced (fold change, 1.8 and 1.3, respectively; $P < 0.01$). This operon structure and the domains of these proteins suggest that they are pseudopilins involved in the biogenesis of a type IV pilus, a required appendage of transformable bacteria. The remaining four genes are of unknown function (*lpp0851*) or belong to the transcription units of the induced genes *radC* and *comM* (*lpp2554*, *lpp2555*/*lepB*, and *lpp0639*). Most likely due to the premature stop codon present in the $lpp0148_{TAA}$ mutant, the operon formed by *lpp0148* and *lpp0149* appears down-regulated (fold change < –2, $P < 0.01$). The function of *lpp0149* is unknown, but it is not needed to repress competence (Fig. S1). The only other down-regulated gene (*lpp0712*) is expressed at a very low level, specific to the Paris strain, and of unknown function. The data show that the constitutive natural transformability of the $lpp0148_{TAA}$ mutant (Fig. S1) is not part of a pleiotropic effect. Rather, Lpp0148 is a specific repressor of a small competence regulon.

Lpp0148 Binds a Highly Conserved Intergenic sRNA Repressor of Competence. Because Lpp0148 carries an RNA-binding domain, we used an RNA immunoprecipitation technique coupled to deep sequencing (RIP-seq) to identify Lpp0148-bound RNAs. Following a reverse transcription protocol that preserves their 5′- and 3′-end sequences, the Lpp0148-bound RNAs were sequenced on an Illumina platform. Illumina reads mapping on five annotated features, including four sRNA features (*lppnc0692*, *lppnc0344*, *lppnc0319*, and *lppnc0187*), were found significantly enriched in the eluted RNAs from the wild-type strain compared with the $lpp0148_{TAA}$ mutant (Table S2). The *lppnc0692* feature showed the highest enrichment (>200-fold), and most importantly, the normalized count of reads mapping on *lppnc0692* represents 99.98% of all combined normalized read counts (Table S2). The reads mapping on *lppnc0692* correspond to a 66-nt sequence beginning at the previously mapped transcription start site (26) and terminating at a predicted Rho-independent transcription terminator (Fig. S2*A*). The RIP-seq data suggest that Lpp0148 nearly exclusively binds in vivo a 66-nt sRNA expressed from the *lppnc0692* feature (Fig. S2*A*). Interestingly, this enriched 66-nt sequence is present in all sequenced *Legionella* species with a sequence identity of over 95%, indicating a functional constraint (Fig. S3). Indeed, RNA fold predicts that this putative sRNA adopts a stable secondary structure with two strong stem-loop structures (SL1 and SL3) and a weak one (SL2) (Fig. 1*A*). Northern blot analysis confirmed the expression of this 66-nt sRNA in the wild-type strain (Fig. 1*B*). Importantly, deletion of the coding region for this sRNA resulted in an increase in *comEA* expression similar to that observed in the $lpp0148_{TAA}$ mutant (Fig. 1*B*) as well as a dramatically enhanced transformability (Fig. 1*C*). Expression of this 66-nt sRNA from a multicopy plasmid reduces transformability and *comEA* mRNA levels in the sRNA deletion mutant but not in the $lpp0148_{TAA}$ strain (Fig. S2*B*). Altogether the data demonstrate that this sRNA is a competence repressor whose function requires the Lpp0148 protein.

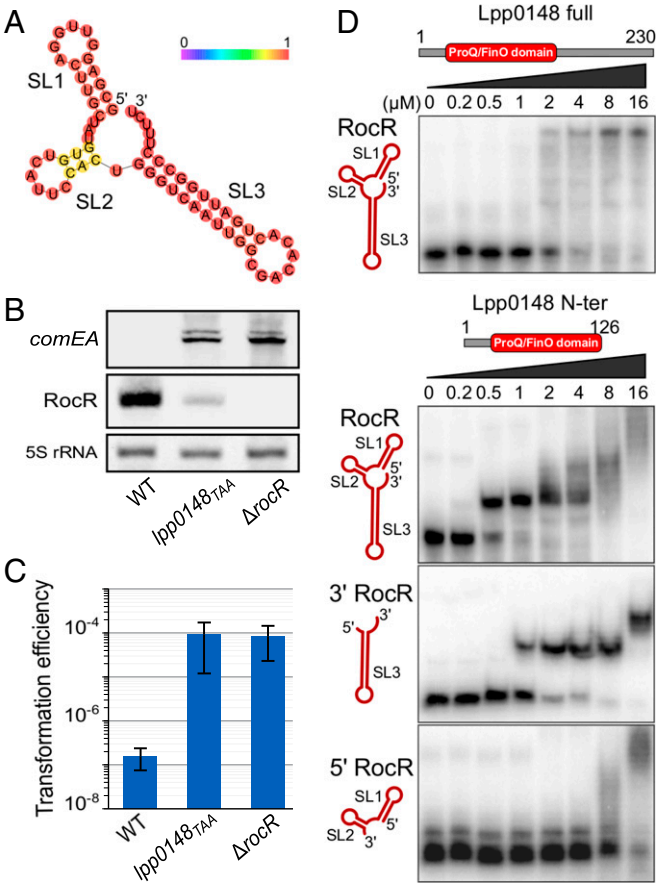

Fig. 1. The intergenic sRNA RocR directly interacts with the ProQ/FinO domain of Lpp0148 and represses competence. (*A*) RNA fold predicted secondary structure of the intergenic 66-nt-long sRNA RocR. Bases are colored according to base-pairing probabilities (see color scale). See also Fig. S2. (*B*) Northern blot analyses of *comEA* and RocR expression in *L. pneumophilla* Paris wild-type, $lpp0148_{TAA}$, and *ΔrocR* strains at an OD_{600} of 0.8 at 37 °C. (*C*) Natural transformability of the same strains as in *B*. Error bars represent SD from the mean of three independent experiments. (*D*) EMSAs of RocR–Lpp0148 complexes. RocR full size, 5′ (SL1 + 2), or 3′ (SL3) parts were incubated with increasing concentrations of Lpp0148 full size or its ProQ/FinO domain in the presence of unlabeled tRNA in excess and run in a native acrylamide gel.

These results prompted us to name it RocR, for "Repressor of competence, RNA."

The ProQ/FinO Domain of Lpp0148 Interacts with the SL3 of RocR. To test the hypothesis that Lpp0148 directly binds RocR, we tested their interaction in vitro by electrophoretic mobility shift assay (EMSA) (Fig. 1*D*). The results indicate that full-length Lpp0148, as well as a truncated protein containing only its ProQ/FinO domain, can bind RocR in the presence of an excess of unlabeled competitor RNA (Fig. 1*D*), suggesting a specific interaction. Moreover, the ProQ/FinO domain of Lpp0148 binds RocR with a higher apparent affinity than full-length Lpp0148 and yields a well-defined shifted band. This suggests a more kinetically stable and homogeneous complex and is reminiscent of FinO, whose proteolytically stable ProQ/FinO domain binds RNA significantly more tightly than the intact protein (20). Our results also indicate that the ProQ/FinO domain of Lpp0148 binds tightly to the 3′ region of RocR (SL3) but only nonspecifically to the 5′ region (SL1 and SL2) (Fig. 1*D*). Taken together, these results indicate that, similarly to FinO (27), Lpp0148 uses its conserved ProQ/FinO domain to directly bind the 3′ rho-independent terminator hairpin and polyU tail of RocR.

The FinO/ProQ Domain of Lpp0148 Protects RocR from Degradation. As a consequence of the hypercompetent phenotype, a plasmid-borne *comEA-gfp* transcriptional fusion is strongly induced in a *lpp0148* deletion mutant and results in bright green colonies (12) (Fig. S4). We used this phenotype to perform a loss-of-function genetic screen. We subjected *lpp0148* to random mutagenesis and isolated 34 *lpp0148* alleles that could not repress competence (Fig. 2*A* and Fig. S4). Among those, we obtained three frameshift mutations leading to premature stop codon and C-terminal truncated proteins (Fig. 2 *A* and *B* and Fig. S4). All other mutants had acquired at least one nonsynonymous mutation in the ProQ/FinO domain, including 10 single mutations that did not alter protein expression (Fig. 2 *A* and *B* and Fig. S4). In these mutants, Northern blot analysis confirmed induction of the chromosomal *comEA* gene to levels similar to those observed in a *Δlpp0148* mutant (Fig. 2*B*). Northern blot analysis also revealed that RocR was less expressed in the $lpp0148_{TAA}$ mutant (Fig. 1*B*) as well as in the *Δlpp0148* mutant and in the strains expressing mutated Lpp0148 proteins than in the wild-type strain (Fig. 2*B*). RNA decay experiments showed that RocR is highly stable with a half-life (~2 h) that exceeds the *L. pneumophila* doubling time in a wild-type strain, but it is relatively unstable in the *Δlpp0148* mutant (Fig. 2*C*). Nonsynonymous mutations in the ProQ/FinO domain that result in the loss of competence repression also resulted in decreased stability of RocR (Fig. 2*C*). Thus, a function of Lpp0148 is to maintain the steady-state level of RocR. However, this may not be sufficient to repress competence, as overexpression of RocR cannot restore competence repression in the $lpp0148_{TAA}$ mutant (Fig. S2*B*). Deletions of the C-terminal domain (through frameshift mutations) did not impact RocR expression and stability to the same extent as the mutations in the ProQ/FinO domain; however, these mutations significantly abrogated repression of competence (Fig. 2 *B* and *C* and Fig. S4). This result is consistent with the finding that the C-terminal domain is not required to bind RocR (Fig. 1*D*) and suggests that this domain could instead be involved in RocR remodeling to promote duplex formation of RocR with its mRNA targets.

Lpp0148 and RocR Target the 5′ UTR of mRNAs Encoding the DNA Uptake System. Lpp0148 interaction with RocR suggests that this ribonucleoprotein complex may directly interact with the mRNAs of the Lpp0148-repressed genes. To identify a possible interaction site, we reexamined the results of the RIP-seq experiment. While specifically looking at Lpp0148-repressed genes (*comEA*, *comEC*, and *comM*) we noticed a sharp peak of coverage at their 5′ UTR (Fig. 3*A*). These sequences were absent or much less abundant in the RNA samples immunoprecipitated

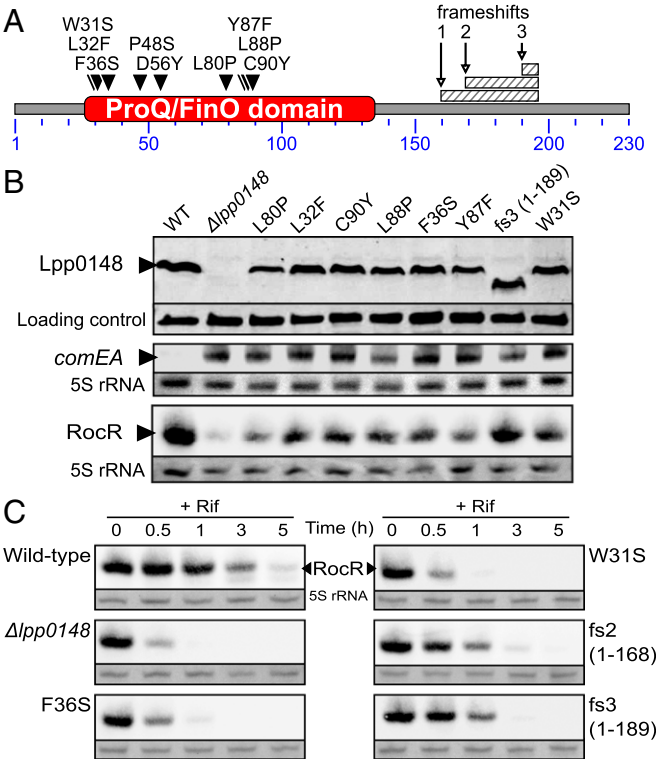

Fig. 2. Mutations of the ProQ/FinO domain of Lpp0148 impair its ability to repress competence and to stabilize RocR. (*A*) Diagram of the *L. pneumophila* Lpp0148 protein. The ProQ/FinO PFAM domain (PF04352) is shown in red. Loss-of-function mutations are indicated as downward-facing black triangles or hatched box for mutations resulting in a frameshift. See also Fig. S4. (*B*) Western blot analysis of Lpp0148 and Northern blot analysis of the competence-induced *comEA* gene in the *L. pneumophila* JR32 wild-type strain and its mutant derivatives (*Δlpp0148* and mutated *lpp0148* alleles). A cross-reacting band and the 5S rRNA were used as loading controls for Western blot and Northern blot, respectively. (*C*) Northern blot analysis of the decay of RocR following transcription inhibition with rifampicin (100 μg/mL) at an OD_{600} of 0.8 in the *L. pneumophila* JR32 wild-type strain and its mutant derivatives.

from the $lpp0148_{TAA}$ and Δ*rocR* mutants, suggesting that they were specifically interacting with Lpp0148 and RocR (Fig. 3*A*). Indeed, RNAfold predicted a potential duplex between the 5′ end region of RocR and the 5′ UTR of Lpp0148-repressed genes (Fig. 3*B*). All predicted duplexes contain a stretch of 7–10 pairing nucleotides near the putative ribosome binding site (RBS) with a shared 6-nt sequence, which we named RocR box, complementary to the first exposed loop of RocR (in SL1). The predicted duplexes are consistent with the prototypical mechanism of sRNA-mediated silencing: binding of the sRNA masks the RBS, thereby preventing translation and exposing the target mRNAs to ribonucleases (28). In agreement with this model, the 1.5-min half-life of the *comEA* mRNA in the wild-type strain is increased to 8–9 min in either the $lpp0148_{TAA}$ or Δ*rocR* mutants (Fig. 3*C*). Substitution of the two consecutive GC pairs of the predicted RocR–*comEA* duplex by two weaker GU wobble pairs resulted in almost complete loss of *comEA* repression (Fig. 3*D*). A double UA pair restored repression, albeit to a lesser extent than the original and stronger GC pairs (Fig. 3*D*). Similarly, changing the following two UA pairs into non-pairing AA affected repression, which was restored by AU pairs. Repression was also affected if the original AGGU sequence of RocR was scrambled into GAUG, but it was restored when the complementary bases were introduced in the *comEA* 5′ UTR (Fig. 3*D*). We conclude that RocR acts by a base-pairing mechanism using its first stem

MICROBIOLOGY

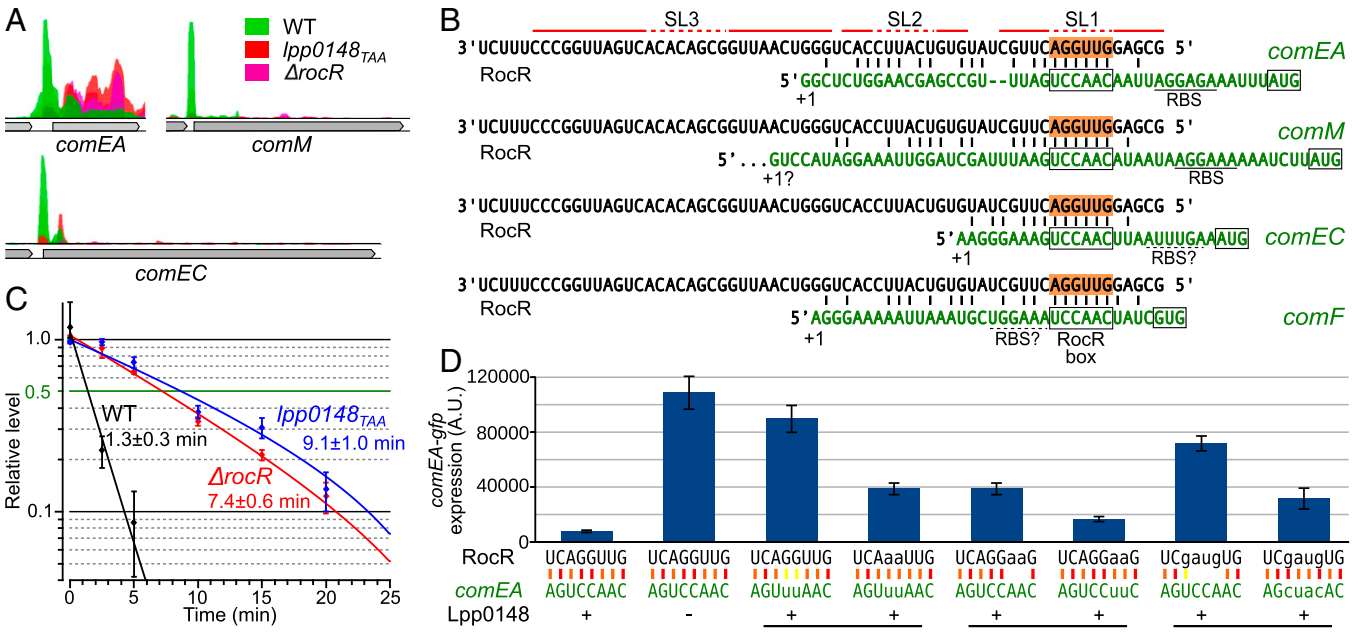

Fig. 3. Lpp0148 and RocR destabilize the *comEA* mRNA by base pairing. (*A*) Strand-specific read coverage of three competence loci obtained by RIP-seq with anti-Lpp0148 antibodies in *L. pneumophila* Paris wild-type (green), *lpp0148*$_{TAA}$ (red), and Δ*rocR* (pink) strains. (*B*) Predicted duplex formation between RocR and the mRNA of Lpp0148-repressed genes. (*C*) *comEA* mRNA half-life determination by RT-qPCR. Decay of *comEA* was followed after transcription was stopped with rifampicin at an OD_{600} of 0.8. Data, expressed as the relative amount of mRNA before the addition of rifampicin (t = 0), were fit to a first-order exponential decay and half-lives were calculated from three quantifications. (*D*) Cultures of different JR32 strains of *L. pneumophila* harboring the plasmid carrying the *comEA-gfp* fusion with wild-type (pXDC91) or mutated (pLLA27-28-29) RocR box were analyzed by flow cytometry. GFP levels were measured in 5.10^5 cells per sample; error bars represent the SD. The strains are JR32 pXDC91 (RocR box WT), JR32 Δ*lpp0148* pXDC91, JR32 pLLA27 (RocR box m3), JR32 *rocRm3* pLLA27, JR32 *rocRm4* pXDC91, JR32 *rocRm4* pLLA28 (RocR box m4), JR32 *rocRm5* pXDC91, and JR32 *rocRm5* pLLA29 (RocR box m5).

loop as a seed sequence to form a duplex with the 5′ UTR of targeted mRNAs and negatively impacts their steady-state levels.

Lpp0148 and RocR Control the Development of Natural Transformability. The *L. pneumophila* Paris strain naturally and transiently develops competence for natural transformation when grown at 30 °C (22). Analyses of the *comEA-gfp* fusion expression (Fig. 4*A*) and chromosomal *comEA* expression by Northern blot (Fig. 4*C*) show that competence (i.e., expression of the DNA uptake system) begins at the midlog phase (OD_{600} > 1.8) and ends before entering the stationary phase (OD_{600} > 4). Expression levels of *comEA* correlate with natural transformability, which is below detection level ($<1 \times 10^{-9}$) in the exponential phase but goes up to a frequency of 2×10^{-5} when *comEA* is expressed, before going back to

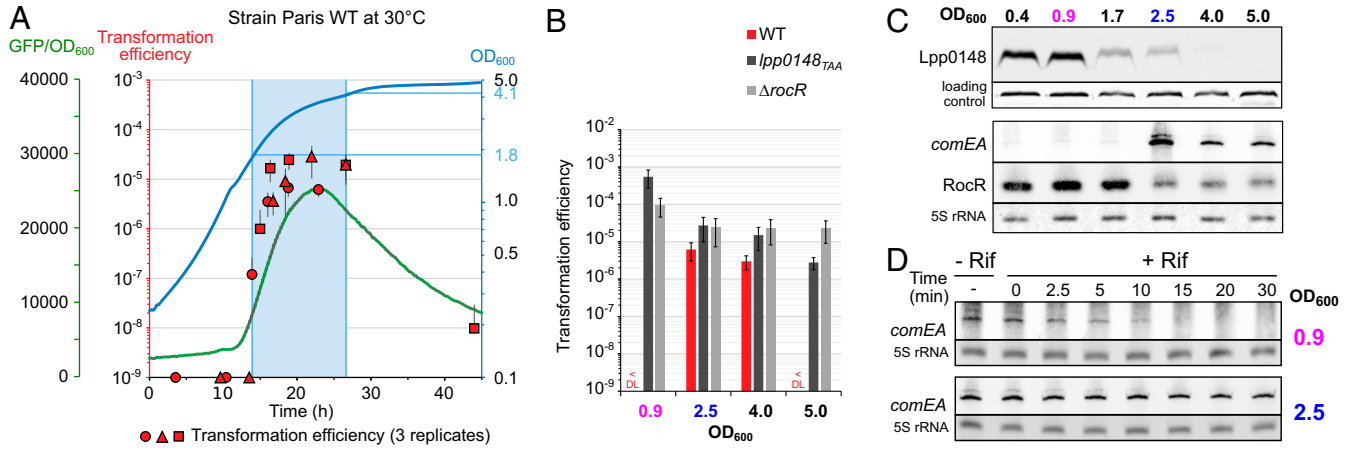

Fig. 4. Lpp0148 and RocR control natural competence development. (*A*) The *L. pneumophila* Paris WT strain transiently develops competence during growth at 30 °C. Expression of *comEA* was followed using a *comEA-gfp* transcriptional fusion carried by pXDC91 (GFP/OD_{600}, green line) and natural transformability (red circle, triangle, and square) determined at different time points during growth (OD_{600}, blue line). Error bars on natural transformation efficiencies data represent SE. (*B*) Natural transformability during growth at 30 °C of the *L. pneumophila* Paris WT, *lpp0148*$_{TAA}$, and Δ*rocR* strains. (*C*) Expression of Lpp0148 and RocR decrease at the onset of the transformability phase at 30 °C. Expression of Lpp0148 was analyzed by Western blot, and expression of *comEA* and RocR was determined by Northern blot analysis. A cross-reacting band and the 5S rRNA were used as loading controls for the Western blot and Northern blot, respectively. (*D*) *comEA* mRNA half-life determination by Northern blot analysis before (OD_{600} of 0.9) and during (OD_{600} of 2.5) the competence phase. Decay of *comEA* was followed after transcription was stopped with rifampicin at the indicated OD. The 5S rRNA was used as a loading control.

basal level in the stationary phase (Fig. 4 *A* and *B*). The observed reduced transformability in the stationary phase is consistent with a previous report that quorum sensing by the Lqs (*Legionella* quorum sensing) system represses competence in the stationary phase (29). In contrast to the wild-type strain, both the $lpp0148_{TAA}$ and Δ*rocR* strains appear highly transformable in all growth phases (Fig. 4*B*). This suggests that a controlled loss of either Lpp0148 or RocR could be responsible for the transient development of competence observed in the wild-type strain. Interestingly, Western blot analysis of Lpp0148 shows that it is expressed in the exponential phase (OD_{600} < 0.9) before its expression steadily decreases at the midlog phase (OD_{600} of 1.7) and at the onset of the stationary phase (OD_{600} of 2.5) to become undetectable in the stationary phase (OD_{600} > 4) (Fig. 4*C*). As Lpp0148 is required to stabilize RocR, its reduced expression could impact the steady-state level of RocR. Northern blot analysis shows that RocR is indeed less expressed starting from an OD_{600} of 2.5 (Fig. 4*C*). The delay between the decrease of expression of RocR compared with that of Lpp0148 is consistent with the high half-life of RocR (Fig. 2*C*). The disappearance of Lpp0148 and RocR correlates with the detection of the *comEA* mRNA (Fig. 4*C*), which becomes highly stable at an OD_{600} of 2.5 (Fig. 4*D*). The data strongly support that competence development in *L. pneumophila* in the midlog phase is triggered by a programmed decreased expression of Lpp0148, which relieves the RocR-mediated silencing of genes encoding the DNA uptake system.

Lpp0148 Belongs to a Diverse Family of ProQ/FinO Domain-Containing Proteins. A systematic survey of 2,775 complete prokaryotic proteomes revealed the existence of 674 distinct proteins containing a ProQ/FinO domain (PF04352) (Dataset S1). Mostly found in Proteobacteria and mostly chromosome-encoded (Fig. S5*C*), they are widespread in Gammaproteobacteria (in 78% of species; Fig. S5*A*), with some species showing up to three or four homologs (Fig. S5*B*). The maximum likelihood phylogenetic tree of ProQ/FinO domain-containing proteins shows two distinct clusters: a compact cluster with short branches and a spread-out cluster with long branches (Fig. 5). Although the former (which includes the *E. coli* ProQ) is consistent with the current taxonomy, the latter presents intermixed sequences from beta- and gammaproteobacteria, which suggests a faster evolutionary rate and an impact of horizontal gene transfer (Fig. 5). This cluster is formed by ProQ/FinO domain-containing proteins with more diverse architecture and includes FinO and Lpp0148 (Fig. 5 and Fig. S5*D*). Broadly, the analysis shows that ProQ/FinO domain-containing proteins form a diverse family of RNA-binding proteins that may control various and unknown processes. Given that the *L. pneumophila* Lpp0148 specifically controls competence development, we propose to rename it RocC, for "Repressor of competence, RNA Chaperone."

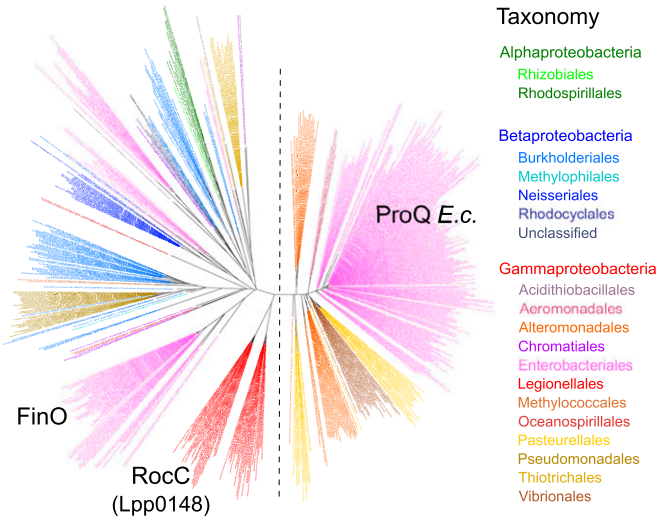

Fig. 5. The ProQ/FinO domain-containing proteins are widespread in Proteobacteria. Shown are the phylogenetic analysis of ProQ/FinO domain-containing proteins (see also Dataset S1 and Fig. S5) and the maximum likelihood phylogenetic tree of the 674 protein sequences containing a ProQ/FinO domain found in 2,775 complete prokaryotic proteomes (80 amino acid positions were used). Colors correspond to taxonomic main lineages.

Discussion

In stark contrast with all known regulatory mechanisms of competence, we report here that competence in *L. pneumophila* is directly controlled by the post-transcriptional repression of genes encoding the DNA uptake system (Fig. S6). This silencing system requires at least two components: a modular *trans*-acting sRNA and an RNA chaperone of non-constitutive expression.

RocR is both the first described sRNA to directly control a competence regulon and the first identified *trans*-acting RNA partner of a ProQ/FinO-domain RNA chaperone. We propose that RocR interacts with the mRNAs of the DNA uptake system encoding genes by a base-pairing mechanism involving a 6-nt sequence, the RocR box, located upstream of the RBS and acting as a seed sequence (Fig. 3 and Fig. S6). Our results suggest that the SL1 and SL2 of RocR form an imperfect duplex with the targeted mRNAs, while the SL3 is engaged in direct interaction with RocC. Similarly to the 3′ end of FinP (27), the RocR SL3 is a hairpin–polyU tail that serves as a rho-independent transcriptional terminator and also specifically binds the ProQ/FinO domain of the protein. Thus, this structure may be a hallmark of sRNAs associated with ProQ/FinO-domain RNA chaperones. Although FinO only works with an antisense (i.e., *cis*-acting) sRNA, RocC interacts with a *trans*-acting sRNA, a function that was, until now, thought to be the exclusive property of the well-studied Hfq protein (28). This functional difference might come from the part of the protein outside the ProQ/FinO domain. Indeed, although this domain alone can bind RNA with higher affinity than the full-length protein (Fig. 1*D* and ref. 20), it is not sufficient to repress conjugation (20) or competence (Fig. 2 and Fig. S4). It was proposed that the flexible N-terminal region of FinO is necessary for sense–antisense RNA pairing (20). Although RocC lacks a similar N-terminal region, it carries an extended C-terminal domain that, despite being dispensable for RocR binding and stabilization, appears required for repression of competence (Fig. 1*D*, Fig. 2, and Fig. S4). We hypothesize that RNA remodeling activities to stimulate base-pairing of RocR with its mRNA targets lie within the C-terminal domain of RocC. Most importantly, RocC exemplifies that the world of RNA chaperones controlling *trans*-acting sRNAs extends beyond the well-documented Hfq protein (30). Chromosome-encoded ProQ/FinO-domain proteins should now be considered bona fide sRNA chaperones that play an important role in the biology of the cell.

We propose here a model in which RocC functions as the master regulator of competence development in *L. pneumophila* by directly controlling expression of the DNA uptake system (Fig. S6). Central to this activity is the regulated expression of RocC, which to our knowledge is the first example of control of a regulon by altering the expression of an RNA chaperone. As the specificity of the targeted regulon lies in a short sequence of the RocC-bound sRNAs, a few base changes may be sufficient to target a different regulon. This represents a way to rapidly evolve a conditional multitarget silencing strategy for disadvantageous processes. As intracellular bacteria thriving on environmental protozoans, *Legionella* species are continuously exposed to the DNA of their defunct eukaryotic hosts and their genome shows numerous acquired eukaryotic-like genes (31, 32). Although some acquired protozoan genes may prove beneficial for *Legionella* to hijack the cellular functions of its host, massive import of foreign DNA may also jeopardize chromosome integrity. The exceptionally conserved RocR sRNA in *Legionella* species (Fig. S3) may have emerged to limit the import of genetic material when

Legionella is exposed to the potentially harmful foreign DNA released by its dead hosts.

Materials and Methods

Detailed protocols for all sections are described in *SI Materials and Methods*.

Bacterial Strains, Plasmids, and Oligonucleotides. See *SI Materials and Methods* and Tables S3 and S4.

Natural Transformability. Transformation ability was determined by incubating bacterial cultures with a non-replicative plasmid containing a kanamycin-resistance cassette that can recombine with the chromosome and produces kanamycin-resistant transformants. Following incubation, serial dilutions were plated on non-selective and selective (i.e., with kanamycin) medium. Transformation frequency is the ratio of the number of cfus counted on selective versus nonselective media.

Gene Expression Analysis and Expression Profiling by RNA-Seq. Total RNA from bacterial cultures was extracted (22), and Northern blot analysis was performed as previously described (12). RNA-seq analysis was performed on DNase-treated RNA samples from bacterial cultures grown to an OD_{600} of 0.8 at 37 °C. Strand-specific cDNA libraries were sequenced on a HiSeq 2000 instrument (ArrayExpress accession no. E-MTAB-4094). Enriched transcripts were determined using DESeq2 (33).

RIP-Seq of RocC-Bound RNAs. Exponentially growing cells were fixed with formaldehyde, lysed, and incubated with protein A magnetic beads coated with affinity-purified antibodies directed against Lpp0148. Bound RNAs were extracted with a tri-reagent solution, and strand-specific cDNA libraries were prepared following a 3′-end polyadenylation and a 5′-end RNA adapter ligation and sequenced on HiSeq 2000.

mRNA Decay and RNA Half-Life Determination. Cultures at an OD_{600} of 0.8 were treated with 100 μg/mL of rifampicin to stop transcription. RNA was extracted at the time points indicated in the figure legends. Transcript levels were analyzed by Northern blot (RocR, *comEA*) and real-time quantitative PCR (RT-qPCR) (*comEA*).

EMSAs. Lpp0148 1–230 and Lpp0148 1–126 were affinity-purified with a 6×His N-terminal tag that was subsequently removed by cleavage with the HRV-3C protease. RNAs were synthesized by in vitro transcription in reactions containing α-[^{32}P]ATP and purified. Binding reactions were carried out for 30 min at room temperature and directly loaded onto polyacrylamide gels run at 4 °C.

Phylogenetic Analysis. A survey of 2,775 complete prokaryotic proteomes was conducted, and the retrieved homologs were aligned. The resulting alignment was trimmed to 80 positions defining the ProQ/FinO domain and used to infer a global phylogeny.

ACKNOWLEDGMENTS. L.A. and X.C. thank Chloé Vallantin for excellent technical support and Christine Gaspin, Unité de Mathématiques et Informatique Appliquées de Toulouse, Toulouse, France, for helpful discussions on sRNA-mRNA duplex predictions. L.A. is supported by a postdoctoral fellowship from the Fondation pour la Recherche Médicale (SPF20130526652). C.B.-A. is a member of the Institut Universitaire de France and is funded by the "Ancestrome" project (ANR-10-BINF-01-01). Work in the A.M.M. laboratory is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Work in the M.G. laboratory is supported by the Canadian Institutes of Health Research (CIHR114975). Work in the X.C. laboratory is supported by a CNRS–INSERM ATIP-Avenir grant and Sanofi. This work was performed within the framework of the LABEX ECOFECT (ANR-11-LABX-0048) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

1. Johnston C, Martin B, Fichant G, Polard P, Claverys J-P (2014) Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* 12(3):181–196.
2. Smith JM, Dowson CG, Spratt BG (1991) Localized sex in bacteria. *Nature* 349(6304): 29–31.
3. Anderson MT, Seifert HS (2011) Opportunity and means: Horizontal gene transfer from the human host to a bacterial pathogen. *MBio* 2(1):e00005–e00011.
4. Overballe-Petersen S, et al. (2013) Bacterial natural transformation by highly fragmented and damaged DNA. *Proc Natl Acad Sci USA* 110(49):19860–19865.
5. Seitz P, Blokesch M (2013) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol Rev* 37(3):336–363.
6. Chen I, Dubnau D (2004) DNA uptake during bacterial transformation. *Nat Rev Microbiol* 2(3):241–249.
7. Provvedi R, Dubnau D (1999) ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Mol Microbiol* 31(1):271–280.
8. Draskovic I, Dubnau D (2005) Biogenesis of a putative channel protein, ComEC, required for DNA uptake: Membrane topology, oligomerization and formation of disulphide bonds. *Mol Microbiol* 55(3):881–896.
9. Mortier-Barrière I, et al. (2007) A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* 130(5):824–836.
10. Attaiech L, et al. (2011) Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: Maintenance of a reservoir for genetic plasticity. *PLoS Genet* 7(6):e1002156.
11. Stone BJ, Kwaik YA (1999) Natural competence for DNA transformation by Legionella pneumophila and its association with expression of type IV pili. *J Bacteriol* 181(5):1395–1402.
12. Charpentier X, Kay E, Schneider D, Shuman HA (2011) Antibiotics and UV radiation induce competence for natural transformation in Legionella pneumophila. *J Bacteriol* 193(5):1114–1121.
13. Buchrieser C, Charpentier X (2013) Induction of competence for natural transformation in Legionella pneumophila and exploitation for mutant construction. *Methods Mol Biol* 954:183–195.
14. Sexton JA, Vogel JP (2004) Regulation of hypercompetence in Legionella pneumophila. *J Bacteriol* 186(12):3814–3825.
15. Kunte HJ, Crane RA, Culham DE, Richmond D, Wood JM (1999) Protein ProQ influences osmotic activation of compatible solute transporter ProP in Escherichia coli K-12. *J Bacteriol* 181(5):1537–1543.
16. Kerr CH, Culham DE, Marom D, Wood JM (2014) Salinity-dependent impacts of ProQ, Prc, and Spr deficiencies on Escherichia coli cell structure. *J Bacteriol* 196(6):1286–1296.
17. Ghetu AF, Gubbins MJ, Frost LS, Glover JN (2000) Crystal structure of the bacterial conjugation repressor finO. *Nat Struct Biol* 7(7):565–569.
18. Arthur DC, et al. (2011) Mapping interactions between the RNA chaperone FinO and its RNA targets. *Nucleic Acids Res* 39(10):4450–4463.
19. Mark Glover JN, et al. (2015) The FinO family of bacterial RNA chaperones. *Plasmid* 78:79–87.
20. Arthur DC, et al. (2003) FinO is an RNA chaperone that facilitates sense-antisense RNA interactions. *EMBO J* 22(23):6346–6355.
21. van Biesen T, Frost LS (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, traJ mRNA, and promotes duplex formation. *Mol Microbiol* 14(3):427–436.
22. Juan P-A, Attaiech L, Charpentier X (2015) Natural transformation occurs independently of the essential actin-like MreB cytoskeleton in Legionella pneumophila. *Sci Rep* 5:16033.
23. Berka RM, et al. (2002) Microarray analysis of the Bacillus subtilis K-state: Genome-wide expression changes dependent on ComK. *Mol Microbiol* 43(5):1331–1345.
24. Peterson SN, et al. (2004) Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays. *Mol Microbiol* 51(4):1051–1070.
25. Redfield RJ, et al. (2005) A novel CRP-dependent regulon controls expression of competence genes in Haemophilus influenzae. *J Mol Biol* 347(4):735–747.
26. Sahr T, et al. (2012) Deep sequencing defines the transcriptional map of L. pneumophila and identifies growth phase-dependent regulated ncRNAs implicated in virulence. *RNA Biol* 9(4):503–519.
27. Jerome LJ, Frost LS (1999) In vitro analysis of the interaction between the FinO protein and FinP antisense RNA of F-like conjugative plasmids. *J Biol Chem* 274(15):10356–10362.
28. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. *Cell* 136(4):615–628.
29. Kessler A, et al. (2013) The Legionella pneumophila orphan sensor kinase LqsT regulates competence and pathogen-host interactions as a component of the LAI-1 circuit. *Environ Microbiol* 15(2):646–662.
30. Gottesman S, Storz G (2015) RNA reflections: Converging on Hfq. *RNA* 21(4):511–512.
31. Cazalet C, et al. (2004) Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity. *Nat Genet* 36(11):1165–1173.
32. de Felipe KS, et al. (2005) Evidence for acquisition of Legionella type IV secretion substrates via interdomain horizontal gene transfer. *J Bacteriol* 187(22):7716–7726.
33. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
34. Charpentier X, Faucher SP, Kalachikov S, Shuman HA (2008) Loss of RNase R induces competence development in Legionella pneumophila. *J Bacteriol* 190(24):8126–8136.
35. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* 97(12):6640–6645.
36. Rutherford K, et al. (2000) Artemis: Sequence visualization and annotation. *Bioinformatics* 16(10):944–945.
37. Law CW, Chen Y, Shi W, Smyth GK (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15(2):R29.
38. Ritchie ME, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7):e47.
39. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402.
40. Eddy SR (2011) Accelerated profile HMM searches. *PLOS Comput Biol* 7(10):e1002195.
41. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30(4):772–780.
42. Criscuolo A, Gribaldo S (2010) BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* 10:210.
43. Price MN, Dehal PS, Arkin AP (2010) FastTree 2–Approximately maximum-likelihood trees for large alignments. *PLoS One* 5(3):e9490.
44. Sadosky AB, Wiater LA, Shuman HA (1993) Identification of Legionella pneumophila genes required for growth within and killing of human macrophages. *Infect Immun* 61(12):5361–5373.