Legionella pneumophila Exploits the CsrA Superfamily To Adapt To a Range of Stresses.

by

Zachary D. Abbott

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in the University of Michigan
2015

Doctoral Committee:

Professor Michele S. Swanson, Chair
Professor Victor J. DiRita
Professor Harry L. T. Mobley
Assistant Professor Kimberley D. Seed

ACKNOWLEDGEMENTS

First and foremost, I must acknowledge my mentor Dr. Michele Swanson. From my first day in her lab, she has been unflinchingly supportive of me in everything I have pursued (even when it took me away from the bench sometimes for months at a time), and she has been an unwavering advocate to all of my scientific and career goals. I have learned so much from Michele, not only about science and research, but also about professionalism, responsibility, prioritization, and personal interaction on all levels. She has been willing to provide as much or as little support and guidance as I have ever needed, and I truly feel that I could not have asked for a better mentor.

I would also like to thank my thesis committee, Drs. Harry Mobley, Vic DiRita, and Kimberley Seed. Harry and Vic have offered support and encouragement both scientifically and personally throughout my tenure at the University of Michigan. Kimberley Seed has graciously stepped up at the last minute to provide her expertise, and I literally could not have defended my thesis without her! In addition, I thank Dr. Blaise Boles, who provided much helpful insight and guidance to this thesis as a committee member before moving to Iowa. Drs. David Friedman, Vern Carruthers, Mary O'Riordan, Joel Swanson, and Phil King have all helped me both professionally and personally, and they have made my experience in the department better.

I am massively grateful to my past and present co-workers: Dr. Zach Dalebroux, Brenda Byrne, Brian Yagi, Karou Harada, Dr. Amrita Joshi, Dr. Maris Fonseca, Dr. Andrew Brian, Kaitlin Flynn, Elisa Hughes, the French-Canadian Sebastien Crepin, Chris Chou, Mahnoor Rehman, and Sara Abrahams. Without you all, the lab would not have been nearly as fun, and I thank you for the wonderful experiences. I would especially like to thank Zach Dalebroux, who mentored me and indeed helped me launch this project with the fateful statement, "there are other Csr's…." Not only was he an excellent mentor, he is a great friend, and I have him to thank for some of my fondest memories in Ann Arbor. Kaitlin Flynn is the little sister I never had, and it's been such an amazing experience getting to work with her these past five years. She is a talented scientist and an amazing woman who will change the world; I have no doubt. Brenda Byrne is truly the greatest human being on the planet, and there are not words to describe how incredible of a scientist and a person she is. Her brilliance at the bench rescued my research at many an impasse, and her bright, positive, and nurturing personality rescued my sanity almost daily.

Finally, my family – Mom and Dad, Scott, Katie, Dan, and Amanda – deserve thanks for all of their support and love throughout this process. It's been a long and sometimes difficult journey, and their unwavering support has sometimes been the only thing that has kept me going.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

The CDC estimates that the bacterium *Legionella pneumophila* causes approximately 8,000-18,000 reported cases of the bacterial pneumonia known as Legionnaire's disease annually in the USA (www.cdc.gov/legionella/clinicians.html). However, humans are a dead end host for the bacteria, because they did not evolve as a human pathogen; there is no human-to-human transmission of this disease, and all human infections are the result of inhalation of contaminated soil or water, the native environment of *L. pneumophila* (Rowbotham 1980). In soil and aquatic environments, the bacteria can persist in biofilms or a metabolically repressed viable but nonculturable (VBNC) form (Steinert et al. 1997; Murga et al. 2001; Declerck et al. 2009; Al-Bana et al. 2014). A diversity of protist species graze upon the bacteria in these states, and *L. pneumophila* has subsequently evolved specifically to survive and replicate within a broad range of protist hosts (Rowbotham 1980; Amaro et al. 2015). For instance, the bacterium can use the *dot/icm* type IV secretion system (T4SS) to manipulate the host and establish a replication niche within the phagocytic vacuole (Qiu et al. 2013; Amaro et al. 2015). Once the niche is established, the bacterium differentiates to a replicative form, characterized by genome-wide changes in gene expression, and replicates within the vacuole to high numbers (Rodgers 1983; Byrne et al. 1998; Faucher et al. 2011;

Weissenmayer et al. 2011). When the bacteria have exhausted the intracellular environment, they again orchestrate a genome-wide expression-profile switch to differentiate into a lysogenic, motile, transmissive form (Byrne et al. 1998; Molofsky et al. 2003; Molofsky et al. 2005; Dalebroux et al. 2009; Edwards et al. 2009). The bacteria lyse out of the host cell where they either encounter another host or persist in the extracellular environment, possibly in a biofilm or the VBNC form (Steinert et al. 1997; Murga et al. 2001; Garduno et al. 2002; Declerck et al. 2009; Al-Bana et al. 2014).

CsrA is a pluripotent regulator

CsrA in the Legionella pneumophila *life cycle*

The genetic switches between the replicative and the trasmissive forms of *L. pneumophila* are well characterized not only phenotypically, but also via microarray and deep sequencing (Faucher et al. 2011; Weissenmayer et al. 2011). In the transition from the replicative to the transmissive form, nutrient stress is detected by cellular sensors that activate a well-defined stringent response (Dalebroux et al. 2009; Edwards et al. 2009; Dalebroux et al. 2010; Dalebroux et al. 2010). At the center of this response is CsrA, a genetic regulator highly conserved in approximately 75% of all bacterial species (Molofsky et al. 2003; Lucchetti-Miganeh et al. 2008; Papenfort et al. 2010). In nutrient rich conditions, *L. pneumophila* requires CsrA to differentiate into the replicative form. As the bacteria replicate, CsrA actively represses translation of proteins associated with transmission and virulence (Molofsky et al. 2003). Specifically, CsrA inhibits motility, pigmentation,

heat and osmotic stress resistance, cytotoxicity, salt sensitivity, and the expression of several *dot/icm* substrates, some of which are known to manipulate host vesicular trafficking. Furthermore, CsrA is essential for *L. pneumophila* to grow in amoebae and macrophages (Molofsky et al. 2003; Forsbach-Birk et al. 2004; Rasis et al. 2009; Nevo et al. 2014).

CsrA in other bacterial species

In addition to its pluripotent functions in *L. pneumophila*, CsrA homologs are key regulators in many other bacterial species (Vakulskas et al. 2015). For example, CsrA regulates glycogen biosynthesis, motility, and biofilm formation in *E. coli* (Wei et al. 2001; Baker et al. 2002; Pannuri et al. 2012); the enterocyte effacement pathogenicity island in EPEC (Bhatt et al. 2009); quorum sensing, motility, and virulence traits in several *Erwinia* species (Mukherjee et al. 1996); swarming motility in *Proteus mirabilis* (Liaw et al. 2003); motility, invasion gene expression, and multiple metabolic pathways in *Salmonella* Typhimurium (Lawhon et al. 2003); motility and adhesion in *Helicobacter pylori* (Kao et al. 2014); hydrogen cyanide production, a type III secretion system (T3SS), a type VI secretion system (T6SS), and iron homeostasis in *Pseudomonas aeruginosa* (Pessi et al. 2001; Brencic et al. 2009); and swarming motility in *B. subtilis* (Mukherjee et al. 2011; Mukherjee et al. 2013). And these are only a few of the many studied CsrA systems. The fact that an exhaustive review of published CsrA functions is impractical in and of itself demonstrates the vast prevalence and diversity of this regulator.

Multiple *csrA* paralogs in a single genome

Despite extensive research on CsrA homologs in dozens of bacterial species, to my knowledge, multiple paralogs within a single genome have been found only in two genera – *Pseudomonas* and *Legionella*. The two *csrA* homologs *rsmE* and *rsmF* (also called *rsmN*) in *Pseudomonas fluorescens* and *Pseudomonas aeruginosa*, respectively, are largely functionally redundant with their paralog *rsmA* (the canonical *csrA* homolog in *Pseudomonas*) (Reimmann et al. 2005; Marden et al. 2013; Morris et al. 2013). Distinct from *Pseudomonas*, the *Legionella* pan-genome encodes several paralogs of *csrA* identified to date. For instance, the original publication of the *L. pneumophila* Philadelphia-1 genome annotated four genes with homology to the canonical *csrA* (Chien et al. 2004). In addition studies of several integrative conjugative elements (ICEs) in the *Legionella* pan-genome noted genes with sequence similarity to *csrA* (Segal et al. 1999; Glockner et al. 2008; Gomez-Valero et al. 2011; Lautner et al. 2013; Wee et al. 2013; Flynn et al. 2014; Gomez-Valero et al. 2014). However, to this point, no published study has analyzed the phylogenetic or functional relationship among the large family of *csrA* paralogs in the *Legionella* genus.

CsrA regulatory mechanism

Direct regulation

CsrA can regulate protein expression by directly binding mRNA targets and affecting their translation. Indeed, in *E. coli* a co-purification of CsrA with bound RNAs demonstrated 721 direct targets (Edwards et al. 2011). Typically, CsrA

binding is repressive. Multiple CsrA proteins recognize GGA motifs in stem-loop secondary structures of mRNAs at the Shine-Dalgarno (SD), prevent protein translation, and mark the mRNA for degradation, as is the case for the target genes *flaB* in *B. burgdorferi* and *hilD* in *S. enterica* (Dubey et al. 2005; Schubert et al. 2007; Martinez et al. 2011; Sze et al. 2011; Lapouge et al. 2013; Duss et al. 2014).

However there are exceptions to this mechanism of repression. For example, the *cel* gene of *E. coli* is an example of CsrA translational repression by binding but without coupled mRNA destabilization (Yang et al. 2010). CsrA represses *sdiA* by binding the mRNA not at the SD, but within the coding region (Yakhnin et al. 2013). Conversely, CsrA binding has occasionally been shown to activate via mRNA stabilization, as with *hrpG* in *Xanthomonas citri*, or via increased translation, as with *moaA* in *E. coli* (Patterson-Fortin et al. 2013; Yakhnin et al. 2013; Andrade et al. 2014). Therefore, in numerous bacterial species, CsrA can directly bind to many different mRNA targets at different positions to differentially repress or activate translation.

Indirect regulation

Direct CsrA-mRNA interactions can have far reaching cellular effects when the mRNA target encodes a transcriptional activator. For example, ~9% of all genes in *P. aeruginosa* are differentially expressed in an *rsmA* mutant, despite the authors inability to detect a direct interaction of RsmA protein with some of the affected mRNAs (Brencic et al. 2009). In *E. coli*, in addition to binding directly to the *pgaABCD* mRNA, which encodes the proteins required for synthesis of the

polysaccharide adhesin PGA and biofilm formation, CsrA prevents the translation of *nhaA*, a transcriptional activator of the *pgaABCD* locus, thereby repressing biofilm formation by a second indirect mechanism (Pannuri et al. 2012). The aforementioned direct repression of *hilD* in *S. enterica* also indirectly represses transcription of two pathogenicity islands, preventing expression of several virulence traits (Martinez et al. 2011). In *E. coli*, by repressing *relA* translation, CsrA indirectly but effectively represses the stringent response (Edwards et al. 2011). Finally, the stabilization of *hrpG* by *X. citri* CsrA ultimately activates the T3SS critical for virulence (Andrade et al. 2014). Therefore, in addition to the direct effects of its binding to mRNA targets, CsrA vastly increases its modulon by regulating genes with downstream regulons of their own.

CsrA structure

Essential to understanding CsrA regulation is knowledge of the structure of the CsrA protein itself and the RNAs it binds. Despite substantial sequence variation among homologues, CsrA protein structure is relatively well conserved. Crystal structures from *E. coli* reveal that CsrA is a homodimer of two *csrA* products, each having five beta-sheets and a short alpha-helix. The homodimer has a hydrophobic core and forms two identical RNA-binding pockets, the sequence of which are highly conserved among CsrA proteins from different species (Gutierrez et al. 2005; Mercante et al. 2006). In particular, the arginine residue at position 44 in the *E. coli* CsrA is ≥ 98% conserved in all surveyed homologues and is essential for RNA binding in CsrA as well as the paralogous proteins from *P. fluorescens* and *P.*

aeruginosa, RsmE and RsmF(N) (Mercante et al. 2006; Schubert et al. 2007; Marden et al. 2013; Morris et al. 2013). I applied knowledge of these conserved binding pocket residues initially to validate the paralogs I identified in the *Legionella* pan-genome.

CsrA regulation

CsrA itself is frequently regulated by two component systems that induce expression of non-coding RNAs that bind and competitively inhibit CsrA. For example, in response to nutrient stress, the *L. pneumophila* two-component system LetA/LetS cooperates with the alarmone ppGpp to activate transcription of two non-coding RNAs, RsmY and RsmZ, which bind to and obstruct CsrA function (Rasis et al. 2009; Sahr et al. 2009). Likewise, in *P. aeruginosa*, the paralogous two-component system GacA/GacS activates transcription of RsmY and RsmZ, and in *E. coli* the paralagous two component system BarA/UvrY activates transcription of the repressive non-coding RNAs, CsrB and CsrC (Heeb et al. 2002; Suzuki et al. 2002; Bordi et al. 2010). However, additional mechanisms also regulate CsrA. For example, in *B. subtilis* the FliW protein binds to CsrA to relieve its repression of flagellin production (Mukherjee et al. 2011). In the dual CsrA system of *P. aeruginosa*, *rsmF (* is directly repressed by RsmA, which binds its mRNA (Marden et al. 2013).

Integrative conjugative elements (ICEs) in the legionellae

ICEs are horizontally acquired mobile genomic elements that direct their own integration into host genomes. These elements encode T4SSs and the machinery that equip the ICE to excise, replicate, and transfer by conjugation to new hosts (Wozniak et al. 2010). ICEs abound in the *Legionella* pan-genome and contribute to the impressive diversity in the *Legionella* genus. It has been speculated that ICEs are responsible at least in part for the legionellae's broad environmental and host range (Miyamoto et al. 2003; Cazalet et al. 2004; Gomez-Valero et al. 2011; O'Connor et al. 2011). One ICE in *L. pneumophila* can complement the entry and intracellular multiplication defects of *dot/icm* mutants (Bandyopadhyay et al. 2007), and another enhances resistance oxidative stress (Flynn et al. 2014).

Interestingly, without exception, *L. pneumophila* T4SS-encoding ICEs contain a *csrA* paralog. The LGI-T4SS ICE lineage characterized by Wee *et al.* all contain a *csrA* paralog (Wee et al. 2013), as do the Tra ICE (Flynn et al. 2014), the Trb ICEs (Glockner et al. 2008; Lautner et al. 2013), and the Lvh ICE (Segal et al. 1999). The *lvrRABC* locus containing the *csrA* paralog *lvrC* was shown to repress excision of the Trb-1 ICE (Lautner et al. 2013). Moreover, canonical CsrA regulates the core *dot/icm* T4SS that is essential for growth of of *L. pneumophila* in amoebae and macrophages (Rasis et al. 2009; Nevo et al. 2014). However, a detailed examination of the phylogenetic and functional relationships among the large family of *L. pneumophila* core and accessory *csrA* paralogs and their related ICEs has not been done.

Outline of thesis

In this thesis, I present the first extensive analysis of the *csrA* gene family in the *Legionella* pan-genome. I identify a second conserved *csrA* paralog in the core genome, as well as several paralogs associated with T4SSs in the accessory genome. Building on the paradigm that canonical CsrA is a central regulator of the *L. pneumophila* replicative-transmissive life cycle, my work examines the central hypothesis that the legionellae exploit the power of their diverse *csrA* superfamily to regulate multiple complex phenotypes that increase the microbe's fitness.

In chapter II, my work is the first to identify and characterize *csrR,* a *csrA* paralog that is part of the core genome and entirely conserved in every surveyed strain in the *Legionella* genus. Previous research identified core *csrA* paralogs in the *P. aeruginosa* and *P. fluorescens* genomes, but in both cases, the paralog was largely redundant with the canonical *rsmA* (Reimmann et al. 2005; Marden et al. 2013; Morris et al. 2013). However in *L. pneumophila*, *csrR* is shown not to be redundant with *csrA*, thus distinguishing it as the first distinct *csrA* paralog. My thesis research also demonstrates that the *csrR* locus enhances *L. pneumophila* survival during prolonged starvation. Although persistence of this pathogen during prolonged incubation in water has been described (Steinert et al. 1997; Al-Bana et al. 2014), my work identifies the first genetic component that contributes to environmental survival. Finally, chapter II also demonstrates *csrR* is a direct target of CsrA post-transcriptional repression, establishing a reciprocal expression model similar to that of *P. aeruginosa rsmA* and *rsmF* (Marden et al. 2013).

In Chapter III, I apply bioinformatics tools to investigate the genetic relationship between *csrA* paralogs and their associated ICEs. Based on my identification of clusters of *csrA* paralogs that appear to be co-inherited with their concomitant T4SS lineages, but not with other associated cargo on their ICEs, I generated the hypothesis that there is a functional relationship between the *csrA* paralogs and their accompanying T4SSs that limits their genetic drift. Extending previous work that characterized one of these ICEs, ICE-βox (Flynn et al. 2014), I show here that ectopic expression of the attendant paralog *csrT* is sufficient to repress ICE-βox conjugation and ICE-βox-mediated oxidative stress resistance. Further, I demonstrate that ICE-βox *csrT* is sufficient to repress two core *L. pneumophila* phenotypes, macrophage infection and motility. Finally, genetic experiments establish that *csrT* can also repress *B. subtilis* motility, supporting the model that ICE-βox CsrT also regulates gene expression by binding mRNA.

Chapter IV I highlights the key contributions of my work. I propose hypotheses and experiments that could expand our understanding of the contribution of *csrA* paralogs in the *Legionella* pan-genome to the diverse life cycle of the legionellae.

CHAPTER II

csrR, a paralog and direct target of CsrA, promotes *Legionella pneumophila* resilience in water

Zachary D. Abbott, Helen Yaknin, Paul Babitzke, and Michele S. Swanson; *mBio* 2015 Vol. 6, No. 3, mBio00595-15

Abstract

Critical to microbial versatility is the capacity to express the cohort of genes that increase fitness in different environments. *Legionella pneumophila* occupies extensive ecological space that includes diverse protists, pond water, engineered water systems, and mammalian lung macrophages. One mechanism that equips this opportunistic pathogen to adapt to fluctuating conditions is a switch between replicative and transmissive cell types that is controlled by the broadly conserved regulatory protein CsrA. A striking feature of legionellae is that each of 14 strains surveyed encodes 4 to 7 *csrA*-like genes, candidate regulators of distinct fitness traits. Here we focus on the one *csrA* paralog (*lpg1593*) that, like canonical *csrA,* is conserved in all 14 strains surveyed. Phenotypic analysis revealed that long-term survival in tap water is promoted by the *lpg1593* locus, which we name *csrR* for CsrA-similar protein for resilience. As predicted by its GGA motif, *csrR* mRNA was bound directly by canonical CsrA protein, as judged by electromobility shift and

RNA-footprinting assays. Furthermore, CsrA repressed translation of *csrR* mRNA *in vivo,* as determined by analysis of *csrR-gfp* reporters, *csrR* mRNA stability in the presence and absence of *csrA* expression, and mutation of the CsrA binding site identified on the *csrR* mRNA. Thus, CsrA not only governs the transition from replication to transmission but also represses translation of its paralog *csrR* when nutrients are available. We propose that, during prolonged starvation, relief of CsrA repression permits CsrR protein to coordinate *L. pneumophila's* switch to a cell type that is resilient in water supplies.

Importance. Persistence of *L. pneumophila* in water systems is a public health risk, yet there is little understanding of the genetic determinants that equip this opportunistic pathogen to adapt and survive in natural or engineered water systems. A potent regulator of this pathogen's intracellular life cycle is CsrA, a protein widely distributed among bacterial species that is understood well. Our finding that every sequenced *L. pneumophila* strain carries several *csrA* paralogs – including two common to all isolates - indicates that the legionellae exploit CsrA regulatory switches for multiple purposes. Our discovery that one paralog, CsrR, is a target of CsrA that enhances survival in water is an important step toward understanding colonization of the human engineered environment by pathogenic *L. pneumophila*.

Introduction

Many gram-negative and gram-positive bacteria rely on a CsrA protein to regulate gene expression (Mercante et al. 2006; Timmermans et al. 2010; Vakulskas et al. 2015). Some of these highly conserved and well-studied RNA-binding proteins repress translation by binding to the 5'-untranslated region (5'-UTR) of target mRNAs and preventing ribosome access to the Shine-Dalgarno (SD). Several species of bacteria practice this mode of post-transcriptional regulation, including *E. coli* (Baker et al. 2002; Dubey et al. 2003; Bhatt et al. 2009; Pannuri et al. 2012), *P. aeruginosa* (Brencic et al. 2009), *B. subtilis* (Yakhnin et al. 2007), *S. enterica* Typhimurium (Jonas et al. 2010), and *L. pneumophila* (Nevo et al. 2014). CsrA binding can also regulate mRNA translation or stability by other mechanisms (Vakulskas et al. 2015). For example, the *E. coli* CsrA protein binds in the 5' coding region of the *sdiA* transcript to repress its translation (Yakhnin et al. 2011).

In the gram-negative, aquatic bacterium *L. pneumophila*, CsrA regulates a key intracellular life cycle switch within host amoeba or mammalian macrophages (Molofsky et al. 2003; Forsbach-Birk et al. 2004; Ohno et al. 2008). *L. pneumophila* differentiates between a replicative form–capable of multiplying to high numbers within a host cell when nutrients are abundant – to a transmissive form – equipped to lyse out of and infect new host cells (Byrne et al. 1998). CsrA is essential for transmissive phase cells to convert to the replicative phase, and it actively represses transmissive phenotypes when nutrients are available to support growth (Molofsky et al. 2003; Forsbach-Birk et al. 2004; Rasis et al. 2009; Nevo et al. 2014). When nutrients become limiting, the LetA/LetS two component system activates transcription of two non-coding RNAs, *rsmY* and *rsmZ*, that directly bind and

sequester CsrA, thus relieving its repression of transmission traits, including expression of *dot/icm* substrates (Hammer et al. 2002; Edwards et al. 2009; Rasis et al. 2009; Sahr et al. 2009; Nevo et al. 2014).

L. pneumophila can also persist in diverse fresh water reservoirs, ranging from ponds to engineered water systems and water-cooling towers (Rowbotham 1980; Fields 1996; Cohn et al. 2014; Osawa et al. 2014; Qin et al. 2014). When machinery vaporizes water contaminated with *L. pneumophila*, bacteria can gain access to the human lung. Remarkably, strategies that evolved to promote persistence within phagocytic amoebae also equip the pathogen to survive and replicate in alveolar macrophages (Horwitz et al. 1980; Fields 1996). Therefore, understanding environmental survival and persistence is critical to designing strategies to reduce the risks *L. pneumophila* poses to human health.

Although distinct extracellular forms of *L. pneumophila* have been described, including biofilm communities, mature intracellular forms (MIFs), and viable but nonculturable cells (VBNCs) (Steinert et al. 1997; Murga et al. 2001; Garduno et al. 2002; Declerck et al. 2009; Al-Bana et al. 2014), the mechanisms that dictate the switch into these environmental forms are not yet known. By applying phenotypic, biochemical and molecular assays, we identify in the *L. pneumophila* core genome a dual CsrA system whose integrated design is predicted to promote reciprocal expression of distinct cell types.

Results

Identification of 71 *csrA*-like genes in *L. pneumophila* species. Initial genome sequencing projects revealed three *csrA*-like genes in the *c*arbon *s*torage *r*egulator family (Chien et al. 2004). By performing a bioinformatic search of the *L. pneumophila* Philadelphia-1 genome, we identified five *csrA*-like genes, including the well-characterized canonical *csrA* (Fig. 2.1A)(Molofsky et al. 2003). The amino acid sequences of the four paralogs exhibit 28% - 58% identity and 59% - 80% similarity.

The amino acid residues determined to be critical for CsrA function in *E. coli* are retained in CsrA proteins across myriad genera surveyed, including the canonical *L. pneumophila* CsrA (Mercante et al. 2006). By aligning the amino acid sequences of the *csrA* paralogs of *L. pneumophila* Philadelphia-1, we confirmed that several key residues are retained in all five paralogs, including the arginine residue that is essential for the RNA-binding activity (Mercante et al. 2006) (Fig. 2.1A). Based on their conserved amino acid sequence, the five *L. pneumophila* Philadelphia-1 loci are likely valid *csrA* paralogs.

To determine if the Philadelphia-1 strain is unique in encoding multiple *csrA* paralogs, the *in silico* search was expanded to include 13 additional fully sequenced legionellae strains: eight additional independent isolates of *L. pneumophila*, two laboratory strains derived from Philadelphia-1, two strains of *L. longbeachiae* and one of *L. drancourtii* (Table 2.1). Altogether 71 *csrA*-like genes were identified, and each retained several residues predicted to be key for RNA-binding (Mercante et al. 2006). In particular, the arginine residue critical for RNA binding was conserved in 70 of the 71 proteins; in the one exception, in place of the arginine was another

cationic residue, lysine (Table 2.2). Each of the 14 surveyed strains contained between four and seven *csrA*-like genes. Without exception, all 14 strains encoded not only canonical *csrA*, but also a copy of a previously uncharacterized paralog, *lpg1593*, which we name here *csrR*. Accordingly, both the *csrA* and *csrR* genes appear to belong to the core genome of *Legionella*. In stark contrast, the remainder of the *csrA*-like genes are located within previously characterized or putative integrative conjugative elements (ICEs), which are large genomic islands known or predicted to transfer horizontally between strains of legionellae (Table 2.3) (Doleans-Jordheim et al. 2006; Wee et al. 2013; Flynn et al. 2014).

To compare the sequences of the 63 legionellae *csrA* genes found in the 12 strains excluding the two laboratory derived strains Lp02 and JR32, we generated a heat map that displays each protein's amino acid similarity to each of the other paralogs. By this approach, we recognized three clusters: two correspond to the presumptive core CsrA and CsrR proteins, and a third encompasses a more diverse group of ICE-related CsrA-like proteins (Fig. 2.1B). Because *csrR* is encoded by all surveyed strains, is highly conserved, and retains key RNA-binding residues, we hypothesized that this gene would regulate traits fundamental to *L. pneumophila*'s life style, as CsrA does.

***L. pneumophila* requires the conserved *csrR* locus to persist in nutrient-poor conditions.** Since *csrR* is a paralog of the essential regulator *csrA*, we first tested whether the two genes regulate the same developmental stage. To do so, we either deleted or induced expression of the *csrR* gene, and then assayed phenotypes known to be regulated by canonical *csrA* (Molofsky et al. 2003). When analyzed in standard

laboratory culture conditions (37˚C in rich AYET medium), neither loss nor gain of *csrR* function altered expression of any *L. pneumophila* replicative or transmissive phase traits tested (data not shown). These pilot experiments included, for broth cultures, growth, pigmentation, motility, and sensitivity to sodium, heat and osmotic stress, as well as infection and intracellular growth in primary mouse macrophages and amoebae (Molofsky et al. 2003). Since *csrR* appeared unlikely to regulate traits that are hallmarks of the commonly studied replicative and transmissive cell types, we next considered conditions that *L. pneumophila* likely encounters in the environment. Indeed, a dual CsrA system (RsmA/RsmF) equips *P. aeruginosa* to modulate biofilm formation (Marden et al. 2013).

The legionellae naturally thrive in fresh water and soil habitats, and *L. pneumophila* can survive for long periods in water in a VBNC state (Steinert et al. 1997; Al-Bana et al. 2014). Therefore, we tested whether *csrR* coordinates differentiation to a resilient environmental cell type. Employing the method of Garduño and colleagues (Al-Bana et al. 2014), we inoculated autoclaved tap water at 45˚C with either wild-type *L. pneumophila*, a *csrR* complete deletion mutant, or a *csrR* mutant strain in which a wild-type *csrR* allele had been integrated into the chromosome. After 24 days, we quantified bacterial viability by fluorescence microscopy using LIVE/DEAD BacLight stain and culturability on rich nutrient agar (Giao et al. 2009). The wild-type culture retained 79% viability, which was significantly higher than the mutant culture's 50% viability (p value = 0.049; Fig. 2.2A). Likewise, the CFU recovered at 24 days relative to day 0 was significantly higher for the wild-type than the mutant culture (5.6% of WT vs. 0.3% of mutant; p

value = 0.0027; Fig. 2.2B). Although the *csrR* merodiploid strain exhibited an intermediate level of viability (67%) as measured by microscopy assay (Fig. 2.2A; p > .05 when compared either to wild type or mutant), genetic complementation was incomplete. Two control experiments indicated that the lack of full complementation is likely due to poor expression of *csrR*. First, qRT-PCR revealed that the *csrR* merodiploid strain contained 2.7% of wild-type levels of *csrR* RNA (Fig. 2.2C). Second, we quantified reduced viability relative to wild-type *L. pneumophila* by both microscopy and CFU assays after isolating and analyzing nine independent *csrR* mutant strains constructed by three different strategies.

***csrR* mRNA contains multiple CsrA binding sites.** Because knowledge of a gene's regulation can provide insight to its function, we next analyzed the upstream sequences of *csrR*. Bioinformatic analysis identified motifs typical of CsrA binding sites in the 5' non-coding region of *csrR* (Nevo et al. 2014). As a first step to analyze *csrR* regulation, we mapped by 5'-RACE its transcriptional start site to –52 relative to the translational start (Fig. 2.3C). Three independent sequencing events revealed that the -52 start site is predominant, although rare transcripts started at -53 or -54. Based on the predominant species, we made several predictions about the *csrR* mRNA. Putative -10 and -35 sites as well as the likely SD approximately 10 bases 5' of the translational start codon were identified (Fig. 2.3A). RNA-fold software predicted a secondary structure for the *csrR* transcript in which the SD is positioned in a loop extending from a long stem (Fig. 2.3B), a common binding motif for CsrA protein (Dubey et al. 2005). Accordingly, we postulated that the *csrR* transcript is bound and its stability regulated by CsrA.

To test directly whether CsrA protein binds to *csrR* mRNA, we performed an *in vitro* assay. Purified CsrA protein was incubated with 5' end-labeled *csrR* RNA, and then their binding was analyzed by gel electromobility shift assay. As predicted by the *in silico* analysis, CsrA protein bound tightly to *csrR* RNA, with a calculated dissociation constant of 32 ± 9nM (Fig. 2.4A). This physical interaction was specific, since incubation with an excess of unlabeled *csrR* RNA effectively competed with CsrA protein binding to labeled *csrR* RNA (Fig. 2.4B). In contrast, incubation with even higher concentrations of a heterologous competitor RNA, *E. coli phoB*, did not inhibit CsrA binding to *csrR*. Therefore, CsrA protein binds *csrR* mRNA tightly and specifically *in vitro*.

To verify our observation that CsrA protein binds to the *csrR* transcript at the predicted stem-loop structure surrounding the SD, we performed RNA footprinting experiments. Interestingly, although CsrA did bind at the predicted SD site ("BS2" in Fig. 2.4C), the protein also protected a GGA motif within the first two codons of the *csrR* coding region ("BS3" in Fig. 2.4C). Although binding in the 5' untranslated region of its target gene is typical, in some cases CsrA does bind in coding regions (Yakhnin et al. 2011). The RNA footprinting assay also revealed a potential third binding site (BS1); however, because this third site was only weakly protected, we focused next on the stronger BS2 and BS3 binding sites.

Transcriptional and translational fusions reveal contribution of BS3 to *csrR* regulation. To analyze in live cells whether either or both binding sites BS2 and BS3 contribute to *csrR* post-transcriptional regulation, we constructed a series of GFP reporters (Fig. 2.5A). For the BS2-only transcriptional reporter, *gfp* expression

is driven by a sequence corresponding to approximately 300 bp immediately 5' of the *csrR* coding region that includes BS2. Although this reporter contains the *csrR* SD, its *gfp* SD likely controls translation; thus the BS2-only construct is referred to as a transcriptional reporter. The BS2/BS3 translational reporter includes both the BS2 and BS3 directly 5' and in frame with *gfp* coding sequences, eliminating the *gfp* SD and including the first 5 codons of *csrR*. For the BS2 stem-loop disrupted (BS2-sld)/BS3 reporter, the BS2 stem-loop structure was disrupted on the translational reporter by changing to uracils the six consecutive adenines immediate 5' of the SD, thereby preventing stem formation with the eight uracils 3' of the SD.

L. pneumophila strains carrying each reporter were cultured in broth, and then aliquots were assessed for GFP fluorescence. As a marker for growth phase, we also analyzed a reporter of the *flaA* flagellin gene, which is strongly induced in PE phase (Hammer et al. 1999). The BS2-only reporter generated a high level of fluorescence in E phase (> 12,000 RFU) and then decreased to ~ 5000 RFU in PE phase (Fig. 2.5B), the period when transmission traits are induced (Byrne et al. 1998). In comparison to the BS2-driven expression, in E phase both the BS2/BS3 and the BS2-sld/BS3 translational reporters demonstrated 3- to 6-fold less fluorescence, respectively (Fig. 2.5B). Therefore, the BS3 site appears to destabilize the mRNA and/or inhibit translation of the transcript. In addition, based on the reduced fluorescence when the stem-loop is disrupted, the BS2 site may somehow stabilize the mRNA.

Mutation of the CsrA binding site relieves CsrA repression *in vivo*. To assess whether CsrA protein destablizes the *csrR* transcript *in vivo*, we utilized a previously

characterized conditional *csrA* mutant (Molofsky et al. 2003). In this strain, the chromosomal *csrA* locus is deleted, and a plasmid carries a functional copy of the gene under IPTG-inducible control. Thus, we measured the impact of CsrA protein on *csrR* transcript levels by culturing cells in the presence and absence of IPTG.

To determine whether CsrA alters the stability of *csrR* transcripts, the conditional *csrA* mutant was cultured with or without IPTG for 9 h, which corresponded to E phase, a period when *csrR* promoter activity is high (Fig. 2.5B) (Sahr et al. 2012). Next, transcription was inhibited by treatment with rifampin. RNA was isolated 0 and 15 min later, and then the relative level of *csrR* mRNA was quantified by qRT-PCR. There was an inverse correlation between the presence of CsrA protein and *csrR* mRNA levels (Fig. 2.6A), indicating that CsrA decreased *csrR* transcript stability. A similar pattern has been reported for other CsrA-mRNA interactions (Liu et al. 1995; Wang et al. 2005).

We then tested *in vivo* whether CsrA destabilization of *csrR* transcript depends on its BS3 motif (Fig. 2.4, 2.5). To do so, we performed qRT-PCR on wild-type *L. pneumophila* strains that carried plasmids in which an IPTG-inducible promoter drives either a wild-type or BS3-mutant *csrR* allele that also encodes a 6x-histidine epitope. To ensure active CsrA, RNA was isolated from strains cultured with or without IPTG for 9 h in E phase. Induction of the BS3-mutant generated > 8-fold more *csrR* transcript than its uninduced control, whereas the wild-type allele increased < 2-fold (p value = 0.033; Fig. 2.6B). Thus, in cultures of wild-type *L. pneumophila*, *csrR* RNA is destabilized by a mechanism that requires its BS3 binding site for the CsrA repressor protein.

Finally, we tested the prediction that, by decreasing *csrR* transcript stability, CsrA repressor binding reduces CsrR protein levels. For this purpose, we performed western analysis on the strains that encode his-tagged CsrR on mRNA that either contained or lacked the BS3 RNA binding site for CsrA protein. CsrR protein was detected only in *L. pneumophila* that carried the *csrR* allele that lacked the BS3 binding site for CsrA (Fig. 2.6C). Therefore, CsrR protein expression is inhibited by the BS3 sequence positioned at the start of the *csrR* opening reading frame, most likely by direct binding of *csrR* mRNA by the repressor protein CsrA (Figs. 2.3-2.5).

Discussion

In nature, legionellae must adapt to the stresses of the intracellular environment of professional phagocytes and nutrient-poor aqueous habitats (Horwitz et al. 1980; Fields 1996; Murga et al. 2001; Declerck et al. 2009; Al-Bana et al. 2014; Kim et al. 2015). *L. pneumophila* relies on the RNA-binding protein CsrA to regulate directly and indirectly the transition between replicative and transmissive phases of its pathogenic life cycle (Molofsky et al. 2003; Forsbach-Birk et al. 2004; Rasis et al. 2009; Nevo et al. 2014). Here we demonstrate that *L. pneumophila* survival during aquatic stress is enhanced by the locus encoding the CsrA-like protein CsrR, a broadly conserved factor whose translation is directly repressed by CsrA.

Of the numerous *csrA* paralogs encoded by *Legionella* species, only two are completely conserved in all surveyed strains: canonical *csrA* (Molofsky et al. 2003) and the gene we name here as *csrR* (Fig. 2.1, Table 2.3). Four genetic observations

establish that CsrR is not redundant with CsrA. First, a *csrA* deletion mutant can not replicate despite the presence of chromosomal *csrR* (Molofsky et al. 2003). Second, inducing expression of *csrR* fails to complement the *csrA* growth defect (pilot experiment; data not shown). Third, neither loss nor gain of *csrR* function alters expression of the panel of traits regulated by CsrA (pilot experiments; data not shown). Instead, the *csrR* locus enhances *L. pneumophila* persistence in 45˚C tap water (Fig. 2.2), a condition that induces resilient VBNC cells (Al-Bana et al. 2014). Together, our initial phenotypic analyses predict that CsrR regulates environmental resilience of *L. pneumophila*, whereas CsrA controls the intracellular replication-transmission cycle (Molofsky et al. 2003).

In addition, CsrA represses CsrR protein expression by a post-transcriptional mechanism. CsrA directly binds *csrR* mRNA *in vitro* (Fig. 2.4); *in vivo*, CsrA destabilizes *csrR* mRNA (Fig. 2.5, 2.6A) and prevents CsrR protein expression (Fig. 2.6C). CsrA binding is mediated by a GGA motif in the second codon of *csrR* (Fig. 2.4C), and mutation of this motif relieves post-transcriptional repression (Fig. 2.6B). CsrA's direct repression of *csrR* ensures reciprocal expression of this dual CsrA system, a design to create two mutually exclusive regulons for *L. pneumophila*.

CsrA represses CsrR protein expression by binding directly to its mRNA at BS3, a site comprised of a GGA-motif within the first two codons of the *csrR*-coding region (Fig. 2.3). It is noteworthy that BS3 overlaps the codon for an aspartic acid residue in the second amino acid position, a striking deviation from the ubiquitous leucine residue found not only in 55 of the 57 non-CsrR CsrA paralogs encoded in the *Legionella* pan-genome, but also in every CsrA homolog of over 30 surveyed

species of gram-negative and –positive bacteria (Table 2.2)(Mercante et al. 2006). This leucine-to-aspartic acid replacement is perfectly conserved in all 14 surveyed *csrR* genes, indicating that the BS3 motif is selected for and retained specifically, likely to conserve the CsrA-mediated repression identified here.

CsrA binds at BS3 in the coding region of the *csrR* transcript (Fig. 2.4), an uncommon, though not unprecedented, location (Yakhnin et al. 2011; Nevo et al. 2014). A second binding site, BS2, was predicted bioinformatically and verified by RNA footprinting (Fig. 2.4). Nevertheless, in our reporter assays, BS2 was not necessary for inhibition. Instead, we observed that mutation of the bases surrounding BS2, which likely disrupt the stem-loop formation that is favorable for CsrA binding, actually decreased expression of the reporter (Fig. 2.5), indicating that BS2 (or at least the stem-loop) may have a stabilizing effect. Alternatively, disruption of the stem-loop may inhibit GFP expression by a CsrA-independent mechanism, or it may enhance rather than disrupt binding by CsrA. More detailed biochemical studies can investigate whether CsrA alternately stabilizes or destabilizes the *csrR* transcript depending on whether it binds at BS2 or BS3, respectively. By analogy to the *E. coli* molybdenum cofactor system (Patterson-Fortin et al. 2013), perhaps a conditional riboswitch governs CsrA binding site selection to ensure *csrR* expression during prolonged starvation.

At the amino acid level, CsrR is only 28% identical to canonical CsrA. Despite significant genetic drift, CsrR retains many of the residues necessary for RNA binding by *E. coli* CsrA (Mercante et al. 2006) (Fig. 2.1, Table 2.2). Likewise, among the homologues of CsrA encoded by many different species of bacteria, the RNA

binding pockets are largely conserved, whereas the rest of the coding region varies considerably (Mercante et al. 2006), as do the targets of regulation. For example, CsrA proteins repress motility of *B. subtilis* by binding directly to the flagellar subunit mRNA (*hag*), but induce motility of *E. coli* by binding to and stabilizing the flagellar regulator *flhDC* mRNA (Wei et al. 2001; Molofsky et al. 2003; Mukherjee et al. 2013; Yakhnin et al. 2013). Evidently, *L. pneumophila* CsrR is under selective pressure to retain RNA binding function. The significant drift in the residues outside of the two binding pockets defined by Mercante *et al.* (Mercante et al. 2006) likely confers CsrR specificity for a distinct regulon.

Our discovery that *csrR* is post-transcriptionally repressed by CsrA (Fig. 2.3-2.5) may explain why *csrR* deletion conferred no discernable phenotype in standard laboratory conditions, which are apparently repressive to CsrR. Likewise, inducing transcription of a plasmid-born *csrR* allele in cells cultured in rich media may be inconsequential because *csrR* mRNAs are bound and degraded by CsrA before they can be translated (Fig. 2.3-2.5). However, inducing CsrR protein expression from a *csrR* allele that lacks the CsrA BS3 binding site (Fig. 2.5) was also not sufficient to trigger *L. pneumophila* differentiation into a resilient, non-replicative cell type. In rich AYE broth, bacterial growth, pigmentation, and motility were not affected by ectopic CsrR protein expression (data not shown). It is possible that the mRNA targets of CsrR are not expressed in the luxuriant culture conditions used, negating the effect of accumulated CsrR protein. We speculate that *L. pneumophila* differentiation into an environmental, resilient cell type requires additional factors to be expressed and/or CsrA repression of other determinants to be relieved.

Although the pathway that triggers CsrA derepression of CsrR remains to be discovered, results herein suggest an expanded model of the *Legionella* life cycle (Fig. 2.7). When a phagocytic cell is encountered, the transmissive form of the bacterium infects and establishes a protective vacuole (Byrne et al. 1998; Roy 2002). The repressor CsrA then equips the intracellular bacterium to differentiate into a replicative form and multiply to high numbers (Molofsky et al. 2003). In replicative phase cells, CsrA actively represses not only transmissive traits, but also accumulation of CsrR, a regulator of environmental persistence (Molofsky et al. 2003; Forsbach-Birk et al. 2004; Rasis et al. 2009; Nevo et al. 2014) (Figs. 2.2-2.5). When nutrients are exhausted within the host cell, LetA/LetS induces expression of the non-coding RNAs RsmY/RsmZ to relieve CsrA repression of transmissive traits (Sahr et al. 2009; Dalebroux et al. 2010; Edwards et al. 2010; Sahr et al. 2012; Nevo et al. 2014). At the same time, *csrR* transcription declines (Fig. 2.5). As a consequence, the cells transition back to the motile, infectious transmissive form that can spread efficiently from one phagocytic cell to another. However, if *L. pneumophila* does not readily encounter another host cell and remains in a nutrient-poor extracellular environment for a prolonged period, we propose that, by mechanisms that remain to be discovered, CsrR is activated and promotes the pathogen's long-term survival (Fig. 2.2).

We favor a model in which the mechanism to relieve CsrA repression of transmissive traits is distinct from its derepression of CsrR protein expression. As nutrients are consumed by replicating bacteria, CsrA repression of virulence traits is relieved at least in part by two non-coding RNAs, RsmY and RsmZ (Rasis et al. 2009;

Sahr et al. 2009; Edwards et al. 2010). In contrast, *csrR* transcript – but not protein – levels are elevated in replicative phase, as shown by our reporter assay (Fig. 2.5), qRT-PCR (data not shown), western analysis (Fig. 2.6) and deep sequencing of RNA pools (Sahr et al. 2012). The high activity of the *csrR* promoter in replicative phase indicates *L. pneumophila* is spring-loaded to respond to abrupt stress, an advantage should the bacteria lack the time or resources needed to complete their natural transition into the transmissive phase. We therefore postulate that some dedicated and yet unknown mechanism relieves CsrA repression of CsrR protein expression. For example, when conditions require rapid deployment of CsrR, replicating cells could induce expression of a direct inhibitor of CsrA binding, or the conditions could facilitate a change in the secondary structure of the *csrR* mRNA that promotes its stabilization by binding of CsrA to BS2 over BS3.

Having two functional *csrA*-like genes in their core genome is an unusual feature of the legionellae. We propose that duplication and genetic drift of the gene endowed these environmental intracellular pathogens with a new regulatory protein that retained the mRNA binding function of CsrA but acquired a new cohort of targets. This evolutionary sequence may have bestowed on the legionellae the capacity to alternate between a facultative intracellular life cycle and a resilient state that can withstand a broad range of environmental conditions. Packaged within this new regulatory *csrR* gene is an on/off switch: a CsrA-binding site in its coding region. We propose that this parsimonious design equips *L. pneumophila's* two CsrA paralogs to mediate reciprocal expression of distinct regulons in appropriate conditions.

Materials and Methods

Bacterial strains, culture conditions, and reagents. Strains, plasmids, and primers used in this study are listed in Table S4. *L. pneumophila* Philadelphia-1 laboratory-derived strain Lp02 , a thymidine auxotroph, was cultured at 37˚C in N-(2-acetamido)-2-aminoethanesulfonic acid (ACES; Sigma)-buffered yeast extract (Becton Dickinson) broth (AYE) and on ACES-buffered charcoal (Fisher) yeast extract agar (Becton Dickinson)(CYE) supplemented with 100μg/mL thymidine (Sigma), 400μg/mL cysteine (Fisher) and 135μg/mL ferric nitrate (J.T. Baker) (Feeley et al. 1979). Thymidine was omitted when culturing thymidine prototroph strains. When necessary for antibiotic selection of mutants or plasmids, media were supplemented with kanamycin (Sigma), 10μg/mL; chloramphenicol (Fisher), 5μg/mL; or gentamicin (Gibco), 10μg/mL. Where indicated, gene expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG; goldbio.com) to a final concentration of 200μM. For all experiments, colonies were first inoculated into broth, incubated overnight, diluted to optical density at 600nm (OD_{600}) 0.05-0.2 and then cultured to E phase (OD_{600} 1.0-2.0) or PE phase (OD_{600} 3.7-4.0), as indicated.

***csrA* gene identification and heat map generation.** To identify the *csrA*-like genes in the *Legionella* pan-genome (Table 2.1), the amino acid sequence of CsrA from *L. pneumophila* Philadelphia-1 was submitted to BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The heat map was generated by first entering the % amino acid identity of each protein pair into an Excel spreadsheet

and then assigning a grayscale to the values. The proteins were then clustered manually by similarity. The order of the proteins from top to bottom (and left to right) is listed in Table 2.2.

***csrR* deletion and merodiploid.** A chloramphenicol resistance cassette was amplified from pKD3 (Datsenko et al. 2000) using primers ZA44+45. Flanking homology to ~500bp 5' and 3' of *csrR* was generated from Lp02 genomic DNA using primers ZA37+43 and ZA46+38, respectively. The three PCR products were then joined by SOEing PCR using the outermost primers ZA37+38 to generate the *csrR::camR* allele used to replace the *csrR* gene by natural transformation (Sexton et al. 2004), producing strain MB1369.

To insert a wild-type *csrR* locus into a neutral site on the chromosome of the *csrR* mutant, the *csrR* gene and its native promoter were integrated into the 154-bp intergenic region between *lpg2528* and *lpg2529* (Liu et al. 2008). To do so, we first constructed plasmid pSU-pZLKm. The multi-cloning site of pSU2719 (Martinez et al. 1988) was amplified using primers SC1+2 and cloned into the *SalI* restriction site of pZL790 (Liu et al. 2008) to create vector pZL790-MCS. Next, a kanamycin resistance cassette flanked by FRT sites was amplified from pKD4 (Datsenko et al. 2000) with primers SC3+4 and cloned into the *SacI* site of pZL790-MCS, creating vector pZL790-MCS-Km. Then, the fragment *lpg2528*-MCS-Km-*lpg2529* from pZL790-MCS-Km was amplified with primers SC5+6 and ligated into pSU2719 at the *Nco*1 and *Hin*dIII sites, after the *Hin*dIII site was made blunt with Klenow fragment, producing vector pSU-pZLKm. *csrR* with its native promoter (~ 300bp 5' of the translational start) was amplified from Lp02 genomic DNA with primers ZA49+60 and ligated into the

*Eco*RI site of pSU-pZLKm. This plasmid was then integrated by natural transformation into the 154-bp intergenic region between *lpg2528* and *lpg2529* in *csrR* mutant MB1369 to generate strain MB1372. Merodiploid strains were verified by growth on CYETcam+kan and by PCR.

Restoration of thymidine prototrophy. *L. pneumophila* is naturally a thymidine prototroph, but in 1993 the laboratory strain Lp02 was made a thymidine auxotroph to help identify intracellular growth mutants (Berger et al. 1993). To avoid the potentially confounding phenotype of thymineless death during prolonged starvation conditions, thymidine prototrophy was restored to strains MB110, MB1369, and MB1372 by replacing by natural transformation their mutant *thyA* allele with the wild-type *thyA* allele from *L. pneumophila* Philadelphia-1 encoded on pJB3395 as described (Sexton et al. 2004), generating strains MB1368, MB1370, MB1371, MB1373, and MB1374.

Water incubation and survival. The thymidine prototroph strains MB1368, MB1370, MB1371, MB1373 and MB1374 were analyzed to avoid thymidine-less death during the prolonged incubation. Wild type and two independent isolates of each mutant and *csrR* merodiploid strain were cultured in AYET to E phase, diluted to OD_{600} 0.2, and then cultured to PE phase. The cultures were then washed twice by suspension in 50mL of autoclaved tap water and centrifugation at 4500xg for 15 min. Cells were then resuspended in 10mL of autoclaved tap water in 15mL conical tubes with the caps loosely affixed to allow air exchange and incubated statically in a 45˚C incubator. Viability of duplicate samples was assessed by LIVE/DEAD BacLight (Life Technologies) staining. A 165μL sample of culture was mixed with 165μL of

autoclaved tap water, 1μL of each stain was added, and the samples incubated at RT in the dark for 15 min. 10μL aliquots were mounted on slides pretreated for 5 min with 5μL poly-L-lysine, and 100 bacteria were scored for each replicate of each strain. Culturability of duplicate samples was quantified by plating 10-fold serial dilutions in autoclaved tap water on CYE.

RNA isolation. For all experiments analyzing RNA, 2mL of liquid culture was pelleted by centrifugation at 10,000xg for 5 min at 4˚C, the cells resuspended in 1mL of Trizol reagent (Life Technologies), and RNA isolated following the manufacturer's protocol. Residual DNA was degraded using TURBO DNA-free kit (Life Technologies) following the manufacturer's protocol, and RNA was stored at -20˚C.

5'-RACE and RNA fold prediction. 5'-RACE was performed using the ligation anchored PCR strategy described previously (Troutt et al. 1992). Briefly, RNA was isolated as described, and first-strand cDNA synthesis was performed using SuperScript II reverse transcriptase (Life Technologies), the manufacturer's protocol, and ZA54 as the *csrR* gene-specific primer. The adaptor oligo ZA80, with 5' phosphorylation for the ligation and 3' amino modifier to prevent its ligation, was ligated to cDNA using T4 RNA ligase (NEB) following the manufacturer's protocol. PCR amplification was performed using ZA81+54 with Platinum Taq DNA polymerase (Invitrogen), and ligation into pGEM T-easy (Promega) was performed following manufacturer's protocols. Sequencing from the pGEM plasmid was performed on three independent clones using primer SL3. mRNA fold prediction was done using IDT DNA UNAFold tool (http://www.idtdna.com/UNAFold?),

inputting the first 70bp of the *csrR* mRNA sequence by using the experimentally determined transcriptional start.

CsrA protein purification. *csrA* with a his-tag epitope was amplified from Lp02 genomic DNA using primers ZA33+34, ligated into plasmid p206gent, and transformed into *E. coli* Rosetta(DE3) pLysS. The strain was cultured in terrific broth medium containing 20μg/mL chloramphenicol and 10μg/mL gentamicin at 37˚C to an OD_{600} ~0.5, supplemented with 500μM IPTG, and cultured for 16 h at 22˚C. Cells were then harvested by centrifugation and flash frozen with liquid N_2. Protein was purified using a 5mL Hi-Trap metal affinity column (GE Healthcare) following manufacturer's instructions. Protein was loaded in a His buffer (25mM NaH_2PO_4, 500mM NaCl, 20mM imidazole) and eluted using an imidazole gradient from 20 to 300mM. Protein-containing fractions were then diluted 1:6 in the elution buffer (due to low solubility) and dialyzed overnight at 4˚C in freezing buffer (10mM Tris-HCl pH 7.4, 10mM $MgCl_2$, 100mM KCl, 10% glycerol).

pYH215 construction. A *csrR* fragment (-168 to +103) was amplified from Lp02 genomic DNA using primers HY2+4 and cloned into the *Eco*RI and *Hind*III sites of the pTZ18U polylinker (Stratagene).

Gel mobility shift assay. RNA was synthesized using the MEGAscript kit (Life Technologies) and a PCR fragment containing a T7 promoter derived from primer HY1. A DNA fragment containing *csrR* sequence extending from +3 to + 84 relative to start of transcription was used as template. Gel-purified RNA was 5′-end labeled with [γ-32P]-ATP. RNA suspended in TE buffer was heated to 90 °C for 1 min followed by slow cooling to room temperature. Binding reactions (10 μl) contained

10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 200 ng/µl yeast RNA, 0.2 mg/ml BSA, 7.5% glycerol, 20 mM DTT, 0.1 nM *csrR* RNA, CsrA-H6 (various concentrations), and 0.1 mg/ml xylene cyanol. Competition assay mixtures also contained unlabeled competitor RNA. To allow CsrA-RNA complex formation, reaction mixtures were incubated for 30 min at 37°C. Samples were then fractionated through native 15% polyacrylamide gels using 0.5X TBE running buffer. Radioactive bands were visualized with a phosphorimager and quantified using ImageQuant 5.2 software. The apparent equilibrium binding constant (Kd) of CsrA-csrR RNA interaction was calculated as described previously (Yakhnin et al. 2000).

RNA footprint assay. RNA was synthesized using the MEGAscript kit and a PCR fragment containing a T7 promoter derived from primer HY3. Template was a DNA fragment containing *csrR* sequence extending from -32 to + 84 relative to start of transcription. Binding reactions (10 µl) containing 5 nM *csrR* RNA and the indicated concentrations of CsrA-H6 were otherwise identical to those in the gel shift assay. After the initial binding reaction, 0.075 U RNase T1 (Roche) was added to the reaction mixtures, and incubation continued for 15 min at 37°C. Reactions were terminated by addition of 10 µl of gel loading buffer and placed on ice. Partial alkaline hydrolysis and RNase T1 digestion ladders of each transcript were prepared as described (Bevilacqua et al. 1998). After fractionating samples through standard 6% polyacrylamide sequencing gels, radioactive bands were visualized with a phosphorimager.

Construction and assays of GFP reporters. To construct the *csrR*-GFP BS2-only transcriptional reporter, ~300bp immediately 5' of the *csrR* coding region was

amplified from Lp02 genomic DNA using primers ZA49+50, ligated directly 5' of GFPmut3 in vector pBH6119, a derivative of pJB98 that encodes a promoterless GFPmut3 locus (Hammer et al. 1999), and then electroporated into Lp02 to generate strain MB1375. For the *csrR*-GFP BS2/BS3 translational reporter, the sequence incorporated into the BS2 reporter was extended to include the first 5 codons of *csrR.* After PCR amplification using primers ZA49+73, the fragment was fused directly 5' and in frame with the second codon of GFPmut3 amplified from pBH6119 using primers ZA72+84 via the SOEing PCR strategy using primers ZA49+84. This product was then ligated into pBH6119, replacing the promoterless GFPmut3, and electroporated into Lp02 to generate strain MB1376. To construct the *csrR*-GFP BS2-sld/BS3 translational reporter, the stem-loop structure of BS2 predicted to favor CsrA binding was disrupted by changing the six consecutive adenines immediate 5' of the SD to thymidines in the DNA sequence (resulting in uracils in the RNA). To do so, the mutation was inserted into two overlapping PCR products using p*csrR*-GFP_BS2/BS3 as template and primer pair ZA49+75 and pair ZA74+84. The two products were then joined via SOEing PCR using primers ZA49+84 to generate the BS2 mutant allele, which was then ligated into pBH6119, replacing the promoterless GFPmut3, and electroporated into Lp02 to produce strain MB1377. For fluorescence experiments, colonies of MB1375-77 were cultured in broth to E phase and then diluted to OD_{600} 0.1. Aliquots were collected at the times shown, normalized to OD_{600} 1, and their fluorescence quantified. Strain MB355, containing p*flaA*-GFP, served as a marker of the entry into PE phase, and strain BH006 carrying pBH6119, encoding GFP but no promoter, was the negative control (Hammer et al.

1999). Growth of all strains was similar, as assessed by OD_{600} readings throughout the experiment (data not shown).

Inducible CsrR constructs. Wild-type *csrR* was amplified from Lp02 genomic DNA using primers ZA59+60, and the PCR product ligated into p206cam (Morales et al. 1991). This plasmid was electroporated into Lp02 to generate strain MB1378.

The BS3 mutant allele of *csrR* was constructed by SOEing PCR. Two PCR products were generated. Product 1 started ~500bp 5' of transcriptional start generated by forward primer ZA37 and included a GAT to TTG mutation in the second codon of *csrR* that was inserted using the reverse primer ZA83. Product 2 encodes the *csrR* coding region, includes the GAT to TTG mutation in the second codon that was inserted by the forward primer ZA82, and extends ~500bp 3' of the coding region by using reverse primer ZA38. 1uL of each of these products then served as template in a SOEing PCR reaction using the outermost primers ZA37 and ZA38. Next, this PCR product containing the BS3 mutation was used as the template to amplify a his-tagged version of only the coding region, using primers ZA59 and ZA60. Finally the product was ligated into p206cam. This plasmid was electroporated into Lp02 to generate strain MB1379.

Rifampin and RNA stability experiment. The conditional *csrA* mutant strain MB464 was cultured on CYE supplemented with 200μM IPTG and then inoculated into AYET containing 200μM IPTG. After culture to E phase, cells were collected by centrifugation, resuspended in AYET without IPTG, and then divided between two tubes that contained or lacked 200μM IPTG. These samples were then cultured to OD_{600} ~1 and aliquots collected for RNA isolation. Next, 100μg/mL rifampin was

added and cultures incubated for 15 min before a second aliquot was collected for RNA isolation.

qRT-PCR. cDNA was generated using 1μL of isolated RNA in a 20μL reaction with iScript cDNA synthesis kit from Bio-Rad following manufacturer's protocol. A 1:50 dilution of cDNA served as template for qRT-PCR in 20uL reactions using iQ SYBR Green Supermix (Bio-Rad). Primers ZA53 and ZA54 were used to assess *csrR* mRNA levels, which were normalized to the internal control 16S rRNA amplified using primers SL1 and SL2 and ΔΔCt analysis.

Western analysis. Strains MB1378 and MB1379 were inoculated from colonies and cultured to E phase, diluted to OD_{600} 0.05, and divided into two tubes, containing or lacking 200μM IPTG. After incubation for ~15 h to E phase, cultures were normalized to OD_{600} 10 by centrifugation of aliquots at 13,000 rpm for 2 min, and then cells resuspended in 100μL Laemmli buffer (2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.005% bromophenol blue, 62.5mM Tris-HCL pH 6.8). Samples were then lysed by boiling for 5 min, and debris pelleted at 13,000 rpm for 2 min. Proteins were separated on a 12% mini-PROTEAN TGX precast gel (Bio-Rad); Precision Plus Kaleidoscope Protein Standard ladder (Bio-Rad) was used as a size marker and to verify transfer. To determine relative loading and transfer of each sample, Ponceau S (Fisher) staining was performed on a duplicate membrane. CsrR was detected via the 6X-histidine epitope that was inserted in-frame at its carboxy-terminus using a 1:5000 dilution of anti-his (C-term)-HRP antibody (Novex, Life Technologies) and the SuperSignal West Pico Chemiluminescent Substrate (Thermo

Scientific). A his-tagged derivative of CsrA paralog Lpg2094 was the positive control.

Table 2.1. Bacterial strains examined for *csrA*-like genes

Species	Strain	NCBI Taxonomy ID
L. pneumophila	Philadelphia-1	272624
L. pneumophila	Philadelphia-1 Lp02	1080311
L. pneumophila	Philadelphia-1 JR32	No NCBI TaxID[a]
L. pneumophila	Paris	297246
L. pneumophila	Lens	297245
L. pneumophila	Corby	400673
L. pneumophila	Alcoy	866628
L. pneumophila	Lorraine	1046632
L. pneumophila	HL06041035	1046631
L. pneumophila	LPE509	1312904
L. pneumophila	ATCC 43290	933093
L. longbeachae	D-4968	638315
L. longbeachae	NSW150	661367
L. drancourtii	LLAP12	658187

Legionella strains surveyed for *csrA*-like genes, with NCBI taxonomy identification numbers indicated. [a] (Rao et al. 2013)

Table 2.2. *csrA* paralogs in *Legionellae* pan-genome

Y-axis pos.	Gene	Species	Strain	Amino acid sequence
1 (top)	***lpg1593***	***L. pneumophila***	**Philadelphia-1**	**MDIINLKFEEPLIIRISNTVVKILAFKTQENGNIKFGV EAPRSINIHREEVFHAIKQKETLSTAD**
2	*lpp1551*	*L. pneumophila*	Paris	MDIINLQFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
3	*lpl1432*	*L. pneumophila*	Lens	MDIINLQFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
4	*lpc1020*	*L. pneumophila*	Corby	MDIINLQFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
5	*lpa02311*	*L. pneumophila*	Alcoy	MDIINLQFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
6	*lpo1613*	*L. pneumophila*	Lorraine HL06041035	MDIINLQFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
7	*lpv1731*	*L. pneumophila*		MDIINLKFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLNTAD
8	*LPE509_01603*	*L. pneumophila*	LPE509	MDIINLKFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
9	*lp12_1531*	*L. pneumophila*	ATCC 43290	MDIINLKFEEPLIIRISNTVVKILAFKTQENGNIKFGVE APRSINIHREEVFHAIKQKETLSTAD
10	*llo1813*	*L. longbeachae*	NSW150	MDIISLQFEEPLMIHIGKASVKILAFKTQEPGNIKFGVD APRSVNVHREEIFNAIKQKQLLDTVE
11	*llb3581*	*L. longbeachae*	D-4968	MDIISLQFEEPLMIHIGKASVKILAFKTQEPGNIKFGVD APRSVNVHREEIFNAIKQKQLLDTVE
12	*ldg7862*	*L. drancourtii*	LLAP12	MDIVSLHFEEPLMININDTIVKILAFKTQEHGNIKFGCE APRSVNVHREEIFHAIKQKQLLEMAE
13	***lpg0781***	***L. pneumophila***	**Philadelphia-1**	**MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAP KDVSVHREEIYLRIQQEKESDDSEQAV**
14	*lpp0845*	*L. pneumophila*	Paris	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
15	*lpl0820*	*L. pneumophila*	Lens	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
16	*lpc2513*	*L. pneumophila*	Corby	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
17	*lpa01193*	*L. pneumophila*	Alcoy	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
18	*lpo0860*	*L. pneumophila*	Lorraine HL06041035	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
19	*lpv0910*	*L. pneumophila*		MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
20	*lp12_0801*	*L. pneumophila*	ATCC 43290	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKESDDSEQAV
21	*llo2071*	*L. longbeachae*	NSW150	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKQSDESEETV
22	*llb3340*	*L. longbeachae*	D-4968	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKQSDESEETV
23	*ldg5259*	*L. drancourtii*	LLAP12	MLILTRRIGETLIIGDDVNITVLGVKGNQVRLGINAPKD VSVHREEIYLRIQQEKQSDDSETTV
24	*LPE509_p00046*	*L. pneumophila*	LPE509	MLILTRRIGETVVIGDDVFITILGIKGNQIRLGFDAPDH VSIHRQEIYLKVQEQKKMRLDSGEAVNGNGMLITPLK QTEPHQYTAH
25	*ldg8306*	*L. drancourtii*	LLAP12	MLILTRRISETMIIGDDVNITILGIKGNQVRLGINAPKSI SVYREEIYLKVKAEEKEQGICLINPKHFTENTSQAVN
26	*lpc0169*	*L. pneumophila*	Corby	MLVLTRRVGESVVIHDDVYCTIVGYRDGEVRLAFDAP RSIPVHRDEIQRRIHRARIKDNWFIDKAANKESIVDRL INKFKNSTSAVKA
27	*lpa00223*	*L. pneumophila*	Alcoy	MLVLTRRVGESVVIHDDVYCTIVGYRDGEVRLAFDAP RSIPVHRDEIQRRIHRARIKDNWFIDKAANKESIVDRL INKFKNSTSAVKA
28	*ldg7118*	*L. drancourtii*	LLAP12	MLVLTRRVGESVVIHDDVYCTIVGYRDGEVRLAFDAP KSIPVHRDEIQRRIYRERQKDKWFHDNAPNKESIVDR LISKFKSSVKLSEYHK
29	*lpc2813*	*L. pneumophila*	Corby	MLVLTRRVGESVVISEDIYCTIVGYQNDEVRLAFDAPK SIPIHRDEIQRRIYRDQIKDNKFVDKAANNESIVDRLIN KFKSSASPTNS
30	*lpa00796*	*L. pneumophila*	Alcoy	MLVLTRRVGESVVISEDIYCTIVGYQNDEVRLAFDAPK SIPIHRDEIQRRIYRDQIKDNKFVDKAANNESIVDRLIN

				KFKSSASPTNS
31	*lpo2456*	*L. pneumophila*	Lorraine	MLVLTRKVGESVVISEEVYCTVVGYRDGEVRLAFDAP QSIPVHRDEIQRRIYRERQRDQWFNDSPSNKENIVDR LISKFKNGLKSA
32	*LPE509_01074*	*L. pneumophila*	LPE509	MLVLTRKVGESVVISEEVYCTVVGYRDGEVRLAFDAP QSIPVHRDEIQRRIYRERQKDQWFSDSPSNKESIIDRLI SKFKHGLKSA
33	*llo2874*	*L. longbeachae*	NSW150	MLVLTRKIGESVVISEDIYCTVVGYRDGEVRLGFDAPQ SIPIHRDEIQRRIYRERQKDQSFNDSPPHKKSIVDRLIS KFKHELKSA
34	*plpl0012*	*L. pneumophila*	Lens	MLILERRVGETVAIENKVFCMVLDHQLDGQLKLAFDA PECVPFHRFEIQEGS
35	*plpp0016*	*L. pneumophila*	Paris	MLILTRRVGDTVVIGNEVFCTVLEQQHDGQIKLAFDA PKSIPIHRFEIQKQIMQKIIEGTYTNDVAWNETVIERLT SQFNRVSHWN
36	*lpop0154*	*L. pneumophila*	Lorraine	MLILTRRIGETVLINDDIYITVLGVKGNQVRLGFDAPQ DVIIHRQEIHQKIKKEQSLFLNKPGQWKEARGVQTH
37	*lpg2094*	*L. pneumophila*	Philadelphia-1	MLSLTRRVGESIVIGEDIFITVLCCKGNQVRIGFNAPNS VAIHRYEIYQKIQSEKHDGLADPTKKFCPSMQQSILNH H
38	*LPE509_01011*	*L. pneumophila*	LPE509	MLSLTRRVGESIVIGEDIFITVLCCKGNQVRIGFNAPNS VAIHRYEIYQKIQSEKHDGLADPTKKFCPSMQQSILNH H
39	*lp12_2035*	*L. pneumophila*	ATCC 43290	MLSLTRRVGESIVIGEDIFITVLCCKGNQVRIGFNAPNS VAIHRYEIYQKIQSEKHDGLADPTKKFCPSMQQSILNH H
40	*LPE509_p00063*	*L. pneumophila*	LPE509	MEIRTVAFESELIISLDNNQKVVITPFKTHVPGNFKLGI DAPKQVTVNRQEIYLRKQEQLKCVFRSEMITNSGLI
41	*LPE509_p00069*	*L. pneumophila*	LPE509	MEIRTVAFESELIITLENNQKVVITPFKTHEPGNFKLG VDAPKYVSINRQEIYLRKQEQLKKEKGQTLTDS
42	*ldg8284*	*L. drancourtii*	LLAP12	MLFCSNYFDIKRENIMEIRPVAFESDLIITLEGNQKIVIT PFKTHEPGNFKLGIDAPKHVTINRQEIYLKKQEQLKKE IRQASADS
43	*lpp1074*	*L. pneumophila*	Paris	MLTLIRRIGEAIYIDKGRIKVHLISEKEGLIRLGIEAPKH VDIERKEVFVRKAVAQHEQAQALRNQSGGDDA
44	*lp12_1033*	*L. pneumophila*	ATCC 43290	MLTLIRRIGEAIYIDKGRIKVHLISEKEGLIRLGIEAPKH VDIERKEVFVRKAVAQHEQAQALRNQSGGDDA
45	*lpo2781*	*L. pneumophila*	Lorraine	MLILIRRMGEAIYIDKGRIKVLLISEKEGLIKLGIDAPKH IDVERKEVFIQKAMEQHALAQKLRDKSTESGGNHA
46	*lpg1003*	*L. pneumophila*	Philadelphia-1	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRAPSEI DIDRKEVFIRKYIQKLDQENKSNQG
47	*lpl1036*	*L. pneumophila*	Lens	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRASSEI DIDRKEVFIRKYIQKLDQENKSNQG
48	*lpc2276*	*L. pneumophila*	Corby	MLILDRKIGEEIFINKGKIKITVLYEKNGLIGIGVRAPSEI DIDRKEVFIRKYIQKLDQENKPNQE
49	*lpa01531*	*L. pneumophila*	Alcoy	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRASSEI DIDRKEVFIRKYIQKLDQENKSNQG
50	*lpv1151*	*L. pneumophila*	HL06041035	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRAPSEI DIDRKEVFIRKYIQKLDQENKSNQG
51	*LPE509_02190*	*L. pneumophila*	LPE509	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRAPSEI DIDRKEVFIRKYIQKLDQENKSNQG
52	*lpp2378*	*L. pneumophila*	Paris	MLVLTRKKGEQIVIDKGQIEIHVIYQRRGVVALGIKAPA HIDVDRKEIFLRKQTNPQNNDKEISK
53	*lpc1860*	*L. pneumophila*	Corby	MLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKAPA HIDVDRKEIFLRKQTNPQNTNIEEPK
54	*lpv1819*	*L. pneumophila*	HL06041035	MLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKAPA HIDVDRKEIFLRKQTNPQNTDIEEPK
55	*lp12_2077*	*L. pneumophila*	ATCC 43290	MMLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKA PAHIDVDRKEIFLRKQTNPQNTNIEEPK
56	*ldg5199*	*L. drancourtii*	LLAP12	MLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKAPP HIDVDRKEIFLRKQANPEYNDDKE
57	*llb1171*	*L. longbeachae*	D-4968	MLVLTRRVGEQIFIDKGQIQIKVLFVRNGNIALGIQAPP NVDVDREEIYYLKKEGIIVESN
58	*lpg1257*	*L. pneumophila*	Philadelphia-1	MLVLTRKAGQQILIGKGLIQMKVLKVDDDIISIGIKAPQ HIDIDREEIYLKKLQQEQAESSMQKVAP
59	*lpl0150*	*L. pneumophila*	Lens	MLVLTRKAGQQILIGKGLIQMKVLKVDDDIISIGIKAPQ HIDIDREEIYLKKLQQEQAESSMQKVAP
60	*lpv0167*	*L. pneumophila*	HL06041035	MLVLTRKAGQQILIGKGLIQMKVLKVDDDIISIGIKAPQ HIDIDREEIYLKKLQQEQAESSMQKVAP

61	*lpp0168*	*L. pneumophila*	Paris	MLVLTRKAGQQILIGKGLIQMKVLKVENDIISIGIKAPS HIDIDREEIYFRKLRQEQEAANDAEMAI
62	*llb1138*	*L. longbeachae*	D-4968	MLILSRKIGENVLIDQGTIQIKLLDVKGRYARIGFIAPA GTDIDREEIYIRKKQSRDLNKKAANHEVNN
63 (bottom)	*llo1850*	*L. longbeachae*	NSW150	MLVLERKIGQKVVIDNGAIEVKVLKPHGDMIRLGFKA PQNMDINKEEIYLRKVLQVPEFLKPVVRNHETRRRK

Each gene is listed with its locus tag identifier in the order it is found in the heat map (Fig. 1B). Canonical CsrA (Lpg0780) and CsrR (Lpg1593) are marked in bold.

Table 2.3. Summary of *csrA*-like genes.

Gene[a]	Size[b]	Prevalence[c]	In ICE?[d]
lpg0781[e]	195	14/14	No
lpg1593	198	14/14	No
lpg1257	204	8/14	Yes
ll02874	261	6/14	Yes
lpc_2813	267	6/14	Yes
lpg1003	201	5/14	Yes
lpc_1860	198	5/14	Yes
lpg2094	237	4/14	Yes
lpp1074	232	4/14	Yes
llo1850	222	2/14	Yes
lpe509_p00069	216	2/14	Yes
llb_1171	186	1/14	Yes
llb_1138	210	1/14	Yes
plpp0016	261	1/14	Yes
plpl0012	159	1/14	Yes
lpe509_p00063	228	1/14	Yes
lpo_p0154	225	1/14	Yes
ldg_8306	234	1/14	Yes

[a] Gene name in one selected strain. Most of these have unique gene annotations in each strain. [b] Size in nucleotides of putative translated region. [c] Prevalence in each of the 14 fully sequenced strains of legionellae surveyed, as available via NCBI BLAST. [d] Putative integrative conjugative elements (ICEs) determined by presence of type IVa secretion system genes immediately adjacent to *csrA*-like gene. [e] Canonical *csrA,* as characterized previously (Molofsky et al. 2003)

Table 2.4. Strains, plasmids, and primers

Strains

Strain number	Name	Description	Source
E. coli			
MB1001	DH5α	F-endA1 *hsdR17* (r- m+) *supE44 thi-1 recA1 gyrA* (Nal^r) relA1 Δ(lacZYA-argF)$_{U169}$Φ80dLacZΔM15λpirRK6	Lab collection
MB1359	Rosetta(DE3) pLysS	F^- ompT hsdS$_B$(R_{B^-} m_{B^-}) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)	Lab collection
MB1360	Rosetta(DE3) pLysS p*csrA*		This study
MB782	DH5α pKD3		Lab collection
MB783	DH5α pKD4		Lab collection
MB1346	DH5α pSU-pZLKm		S. Crepin, unpublished
MB1361	DH5α pSU-pZLKm-csrR		This study
MB643	DH5α pJB3395		(Sexton et al. 2004)
MB1362	DH5α pGEM csrR_5'-RACE		This study
BH003	DH5α pBH6119		(Hammer et al. 1999)
MB1363	DH5α pcsrR-GFP_BS2		This study
MB1364	DH5α pcsrR-GFP_BS2/BS3		This study
MB1365	DH5α pcsrR-GFP_BS3		This study
MB1068	DH5α p206cam		(Hammer et al. 1999)
MB1366	DH5α pcsrR-his		This study
MB1367	DH5α pcsrR-his_BS3M		This study
L.pneumophila			
MB110	Lp02	Philadelphia-1 Str^r HsdR− Thy−, wild-type strain for this study	(Berger et al. 1993)
MB1368	Lp02 thy+	MB110 with thymidine autotrophy restored	This study
MB1369	Lp02 ΔcsrR::Cam^r	MB110 with *csrR* replaced with chlormaphenicol resistance cassette	This study
MB1370	Lp02 ΔcsrR::Cam^r thy+ #2	MB1369 with Thymidine autotrophy restored, clone #2	This study
MB1371	Lp02 ΔcsrR::Cam^r thy+ #4	MB1369 with Thymidine autotrophy restored, clone #4	This study
MB1372	Lp02 ΔcsrR::Cam^r + csrR	MB1369 with *csrR* replaced on chromosome between *lpg2528* and *lpg2529;* Cam^r Kan^r	This study
MB1373	Lp02 ΔcsrR::Cam^r + csrR thy+ #1	MB1372 with Thymidine autotrophy restored, clone #1	This study
MB1374	Lp02 ΔcsrR::Cam^r + csrR thy+ #3	MB1372 with Thymidine autotrophy restored, clone #3	This study
BH006	Lp02 pBH6119	MB110 with promoterless GFP vector, negative control for reporter experiments	(Hammer et al. 1999)
MB1375	Lp02 p*csrR*-GFP_BS2	MB110 with *csrR* BS2-only transcriptional GFP reporter	This study
MB1376	Lp02 p*csrR*-GFP_BS2/BS3	MB110 with *csrR* BS2/BS3 translational GFP reporter	This study
MB1377	Lp02 p*csrR*-GFP_BS3	MB110 with *csrR* BS2-sld/BS3 translational GFP reporter	This study
MB355	Lp02 p*flaA*-GFP	MB110 with *flaA* GFP reporter, positive control for reporter experiments	(Hammer et al. 1999)
MB464	Lp02 Δ*csrA*::Kan^r p*csrA*	*csrA* conditional mutant; Kan^r, pMMB206Δmob *csrA,* inducible *csrA*, cam^r	(Molofsky et al.

			2003)
MB1378	Lp02 p*csrR*	MB110 with IPTG-inducible *csrR*	This study
MB1379	Lp02 p*csrR*-BS3M	MB110 with IPTG-inducible *csrR* BS3 mutant	This study

Plasmids

Plasmid	Description	Source
pKD3	Template plasmid for the amplification of the chloramphenicol gene bordered by FRT sites, Amp^r; can	(Datsenko et al. 2000)
pSU2719	Cloning plasmid; P15A replicon, *lacZ*α, Cam^r	(Martinez et al. 1988)
pZL790	Suicide plasmid; *ori* R6K, *sacB*, Tet^r	(Liu et al. 2008)
pZL790-MCS	Suicide plasmid; *ori* R6K, *sacB*, MCS of pSU2719, Tet^r	S. Crepin, unpublished
pKD4	Template plasmid for the amplification of the kanamycin gene bordered by FRT sites, Amp^r; Kan^r	(Datsenko et al. 2000)
pZL790-MCS-Km	Suicide plasmid, *ori* R6K, *sacB*, MCS of pSU2719 + Km, Kan^r; Tet^r	S. Crepin, unpublished
pSU-pZLKm	Cloning plasmid; pSU2719 backbone; backbone vector for chromosomal integration into intergenic re	S. Crepin, unpublished
pSU-pZLKm-*csrR*	plasmid for insertion of *csrR* onto chromosome between *lpg2528* and *lpg2529* by homologous recomb	This study
pJB3395	*L. pneumophila* Philadelphia-1 *thyA* ligated into BamHI/SalII of pBluescript, for restoring thymidine au	(Sexton et al. 2004)
pGEM T-easy	linearized vector with T-overhangs for ligating in Taq-amplified PCR products; Amp^r	Promega
pGEM csrR_5'-RACE	pGEM T-easy with *csrR* cDNA including 5' untranslated region ligated into T-overhangs, Amp^r	This study
p206gent	pMMB66EH derivative, Δ*mob, lacI*^{q}, $P_{taclacUV5}$, $gent^r$	(Dalebroux et al. 2010)
p*csrA*-his	p206gent with *csrA* with 6x-his epitope ligated into the SalI and HindIII sites, IPTG-inducible *csrA*-his;	This study
pJB98	pKB5 plasmid with HpaI/EcoRI fragment containing the Ptac promoter and the lacIq gene removed; a	(Hammer et al. 1999)
pBH6119	XbaI/PstI fragment of GFPmut3 insterted into pJB98; Amp^r, TdΔi	(Hammer et al. 1999)
p*csrR*-GFP_BS2	309bp *csrR* promoter fragment fused to GFPmut3 in pBH6119; Amp^r, tdΔi	This study
p*csrR*-GFP_BS2/BS3	323bp *csrR* promoter and first 5 codons of *csrR* coding region fused in frame to GFPmut3 in pBH6119	This study
p*csrR*-GFP_BS3	p*csrR*-GFP_BS2/BS3 with stem-loop structure of SD disrupted; Amp^r, tdΔi	This study
p*flaA*-GFP	~150bp *flaA* promoter fragment fused to GFPmut3 in pBH6119; Amp^r, tdΔi	(Hammer et al. 1999)
p206cam	pMMB66EH derivative, Δ*mob, lacI*^{q}, $P_{taclacUV5}$, Cam^R	(Morales et al. 1991)
p*csrR*-his	p206cam with *csrR* with 6x-his epitope ligated into the EcoRI and HindIII sites, IPTG-inducible *csrR*; C	This study
p*csrR*-his_BS3M	p206cam with *csrR* BS3 mutant with 6x-his epitope ligated into the EcoRI and HindIII sites, IPTG-indu	This study
pYH215	csrR fragment (-168 to +103) inserted into EcoR1 and HindIII sites of pTZ18U plasmid from Stratager	This study

Primers

Number	description	sequence	Source
ZA37	*csrR* SOEing front flank fwd	CCTTCAGCACTAAAGCGGCAC	This study
ZA43	*csrR* SOEing front flank rev	AGGAACTTCGAAGCAGCTCCAGCCTACAGACCTCTTTTTTCATAAATTA	This

		TTCCTAAGGCC	study
ZA44	*csrR* SOEing resistance fwd	GGCCTTAGGAATAATTTATGAAAAAAGAGGTGTGTAGGCTGGAGCTGC TTCGAAGTTCCT	This study
ZA45	*csrR* SOEing resistance rev	CTGAAAGGTATAACTTGCATTTATTTGTTACATATGAATATCCTCCTTA GTTCCTATTCC	This study
ZA46	*csrR* SOEing back flank fwd	GGAATAGGAACTAAGGAGGATATTCATATGTAACAAATAAATGCAAGT TATACCTTTCAG	This study
ZA38	*csrR* SOEing back flank rev	GTTCAGGGGCGAATAAACTGCTTTG	This study
ZA33	SalI *csrA*-his fwd	GTCGACGGAAGCGTTCTATGGAGAAGTAGG	This study
ZA34	HindIII *csrA*-his rev	AAGCTTTTAATGATGATGATGATGATGTACTGCTTGTTCCGAATCATC AG	This study
ZA49	*csrR* promoter fwd EcoRI	TCATGAATTCTCCTTATGGATAAAGCCTTCCAGC	This study
ZA50	*csrR* BS2-only rev BamHI	GATCGGATCCTTAAAAAAAAACCTCTTTTTTCATAAATTATTCCTAAGG	This study
ZA53	*csrR* qPCR fwd	GAACCACTTATCATTCGTATCTCG	This study
ZA54	*csrR* qPCR rev	GCATGGAAAACCTCTTCTCTG	This study
ZA59	EcoRI *csrR*-his fwd	CATGGAATTCGTACCGTTCAGCAGTGAATATTGG	This study
ZA60	HindIII *csrR*-his rev	CATGTTCGAATTAATGGTGATGGTGATGGTGATCAGCTGTGCTAAGAG TTTC	This study
ZA72	*csrR* BS2/BS3 SOEing fwd	GAGGTTTTTTTTAAATGGATATAATAAATATGAGTAAAGGAGAAGAAC TTTTCACTGG	This study
ZA73	*csrR* BS2/BS3 SOEing rev	CCAGTGAAAAGTTCTTCTCCTTTACTCATATTTATTATATCCATTTAAA AAAAACCTC	This study
ZA74	*csrR* BS2-sld SOEing fwd	GGCCTTAGGAATAATTTATGTTTTTTGAGGTTTTTTTTAAATGG	This study
ZA75	*csrR* BS2-sld SOEing rev	CCATTTAAAAAAAACCTCAAAAAACATAAATTATTCCTAAGGCC	This study
ZA80	RACE anchor	/5Phos/TTTAGTGAGGGTTAATAAGCGGCCGCGTCGTGACTGGGAGCG C/3AmMO/	This study; (Troutt et al. 1992)
ZA81	RACE PCR fwd primer	GCGGCCGCTTATTAACCCTCACTAAA	This study; (Troutt et al. 1992)
ZA82	*csrR* BS3 mut SOEing fwd	GAGGTTTTTTTTAAATGTTGATAATAAATCTTAAATTTG	This study
ZA83	*csrR* BS3 mut SOEing rev	CAAATTTAAGATTTATTATCAACATTTAAAAAAAACCTC	This study
ZA84	sphI GFP SOE rev	CCAAGCTTGCATGCCTG	This study
SL1	16S fwd	AGAAGGCCTGAGGGTTGTAAAGCA	lab collection
SL2	16S rev	ACCCTTTACGCCCAGTAATTCCGA	lab collection
SL3	pGEM fwd for sequencing	GATTAAGTTGGGTAACGCCAGGGT	lab collection
SC1	MCS_pSU2719_F_Xho1	TGTTTTTGCTCGAGTCCAGTGGCTTCTGTTTCTATC	This study
SC2	MCS_pSU2719_R_Xho1	GTGAGCGGCTCGAGTTCACACAGGAAACAGCTATGA	This study
SC3	Km_Cm_F_Sac1	GTCTTGAGGAGCTCTGTAGGCTGGAGCTGCTTCG	This study
SC4	Km_Cm_R_Sac1	GTAACGCAGAGCTCAGCCCTTAGAGCCTCTCAAAGCAA	This study
SC5	Clon_pZLfrag_Km_Cm_F_Pci1	ATGAACATGTTGATAACCCAAGAGGGCATTT	This study
SC6	Clon_pZLFrag_Km_Cm_Gm_R_Stu1	CCGGAGGCCTACGCCATTCATGGCCATATC	This study
SC7	lpg2528_pZL_F	AACGCAAATTTCAACATGCCCCGC	This study
SC8	lpg2529_pZL_R	ATGGAGCAAGTCCGTCCTATCAG	This study
HY1	GelSH fwd T7CsrR	GAAATAATACGACTCACTATAGGGAGTGAATATTGGCCTTAGGAATAA	This study
HY2	csrR rev HindIII	CCCAAGCTTCGAGATACGAATGATAAGTGGTTC	This study
HY3	Footprint fwd T7	GAAATAATACGACTCACTATAGGGTGACAACCGTCACCAAAGTAAG	This study
HY4	csrR fwd EcoRI	CGGGAATTCGTCCTTACAATCTCATTGCGA	This

HY5	csrR rev	GGTTCTTCAAATTTAAGATTTATTATATCC	This study

Strains, plasmids, and primers used in this study.

Figure 2.1

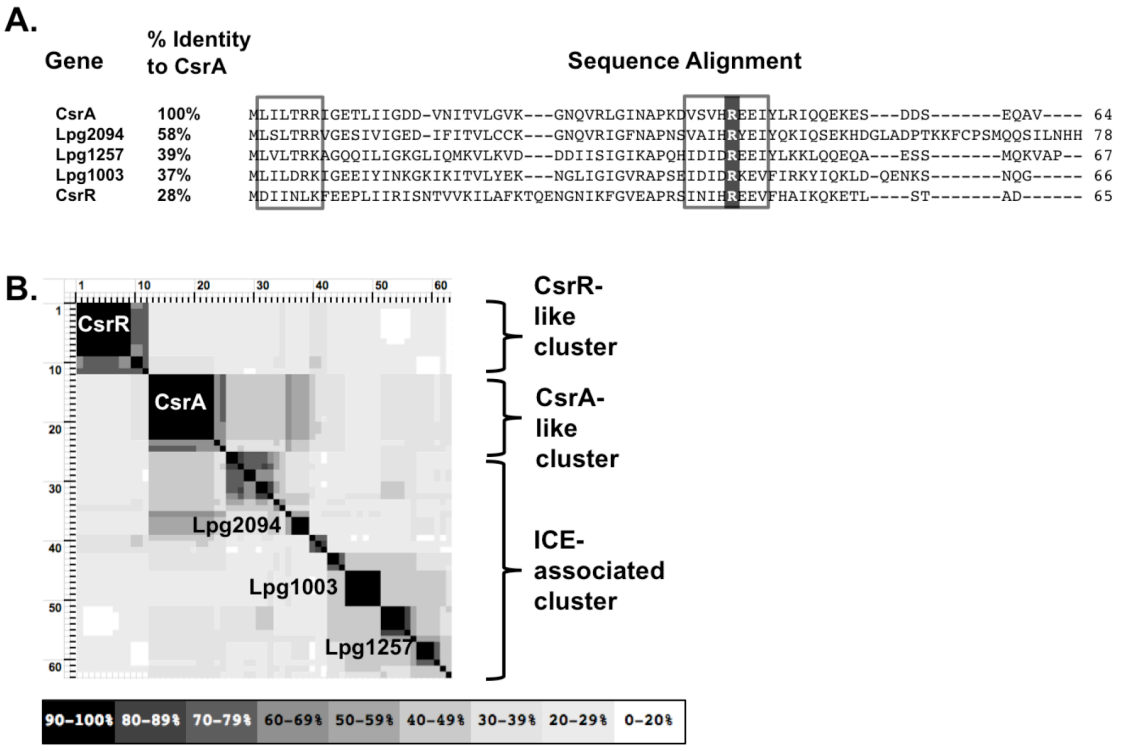

Figure 2.1. CsrA paralogs in 12 strains of legionellae identified *in silico*.
(A) Alignment of the five CsrA-like proteins in *L. pneumophila* Philadelphia-1 with the essential arginine residue highlighted and the two RNA binding pockets outlined (Mercante et al. 2006). (B) Sequence identity of the 63 CsrA-like proteins found in 12 legionellae strains. Darker shading indicates greater % identity. Each of the five CsrA proteins encoded by *L. pneumophila* Philadelphia-1 is marked, and clusters are named for two core and the remaining ICE-associated genes.

Figure 2.2

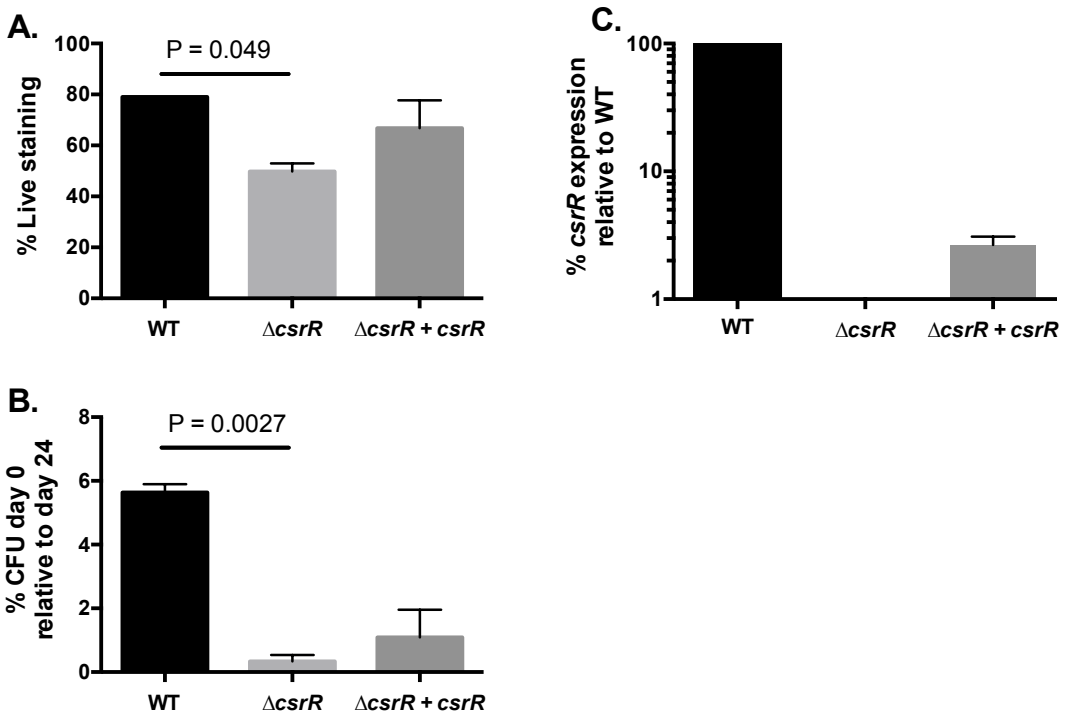

Figure 2.2. *csrR* enhances survival of *L. pneumophila* in water. After incubation for 24 days in 45°C tap water, survival of wild type (MB1368, black), Δ*csrR* deletion mutants (MB1370 and MB1371, light grey), and genetically complemented strains strains (MB1373 and MB1374, dark grey) was assessed (A) microscopically by LIVE/DEAD fluorescence staining and (B) by enumerating CFU. Error bars represent standard deviation calculated from duplicate samples from two independent isolates of each genotype analyzed in parallel. Significance was calculated using a student's *t*-test. Differences reported for the wild-type and Δ*csrR* strains are representative of multiple independent experiments. (C) Expression of *csrR* transcript in the strains listed above relative to wild type expression was quantified by qRT-PCR. Error bars represent standard deviation of duplicate wells from a single experiment. Data is representative of two independent experiments.

Figure 2.3

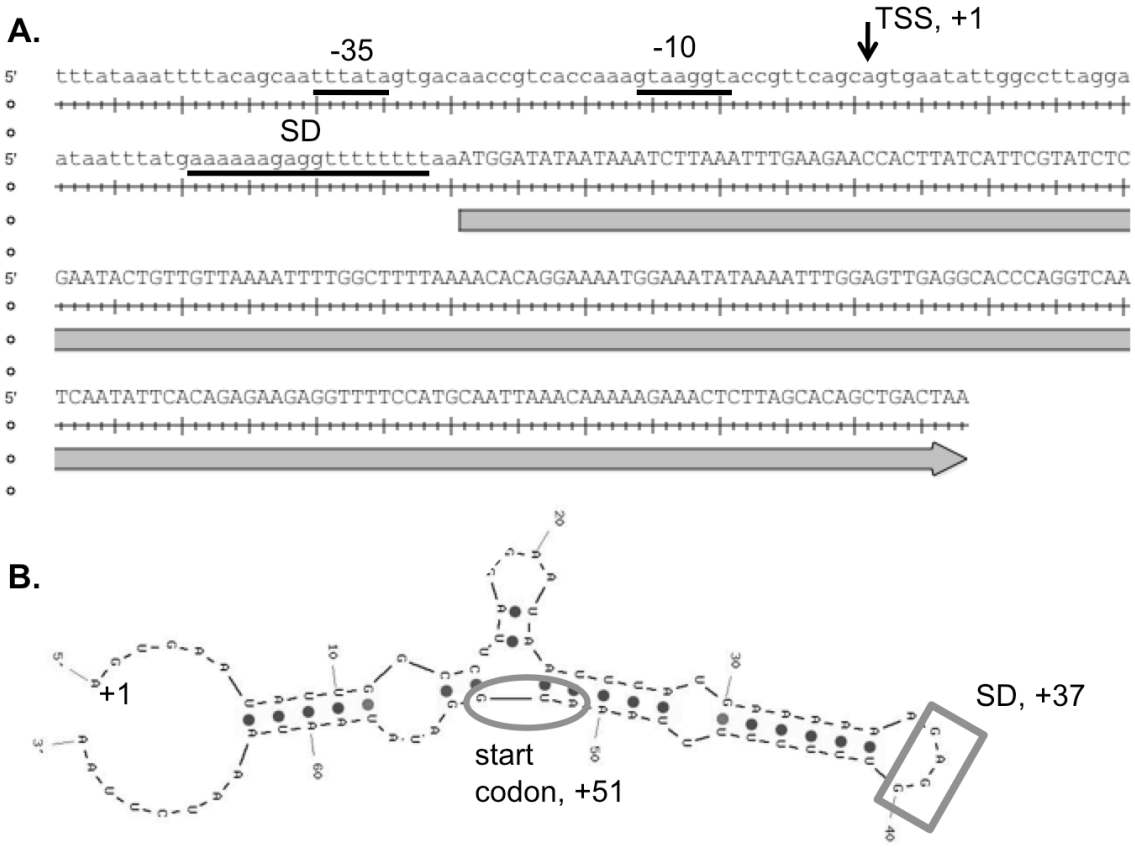
A.
TSS, +1
-35
-10
SD
5' tttataaattttacagcaatttatagtgacaaccgtcaccaaagtaaggtaccgttcagcagtgaatattggccttagga
5' ataatttatgaaaaaagaggttttttttaaATGGATATAATAAATCTTAAATTTGAAGAACCACTTATCATTCGTATCTC
5' GAATACTGTTGTTAAAATTTTGGCTTTTAAAACACAGGAAAATGGAAATATAAAATTTGGAGTTGAGGCACCCAGGTCAA
5' TCAATATTCACAGAGAAGAGGTTTTCCATGCAATTAAACAAAAAGAAACTCTTAGCACAGCTGACTAA
B.
+1
start
codon, +51
SD, +37
C.
Reverse complement
of csrR 5' UTR
Reverse complement
of ZA81 primer
TAAGGCCAATATTCACTTTTNG

Figure 2.3. A putative CsrA binding site on the *csrR* RNA overlaps with a ribosome binding site motif. (A) The *csrR* transcriptional start site (TSS, +1, black arrow) was mapped by 5'RACE. Likely -10 and -35 promoter elements are underlined and labeled; a putative Shine-Dalgarno (SD) and the associated flanking bases predicted to form a stem-loop structure recognized by CsrA and located 10 bases 5' of the translational start are underlined; and the *csrR* open reading frame is marked with the green arrow and capitalized letters. (B) The *csrR* RNA encodes a putative SD (+37) located on a stem-loop and an ATG start codon (+51) positioned at the base of two stem-loops, as predicted by the IDT UNAFold algorithm. CsrA protein is predicted to bind the *csrR* RNA SD (boxed). (C) The chromatogram of the sequencing reaction from the 5'-RACE experiment. *csrR* transcript fused to primer ZA80 was sequenced using ZA54 and ZA81. The string of three thymidines mark the beginning of the ZA80 primer, demonstrating that transcription is most often initiated at the adenine at position -52 relative to the translational start codon. However small peaks beneath the first two thymidines indicate potential minority transcript species commencing at the -53 and -54 positions. The chromatogram is representative of sequencing reactions from three independent isolates.

Figure 2.4

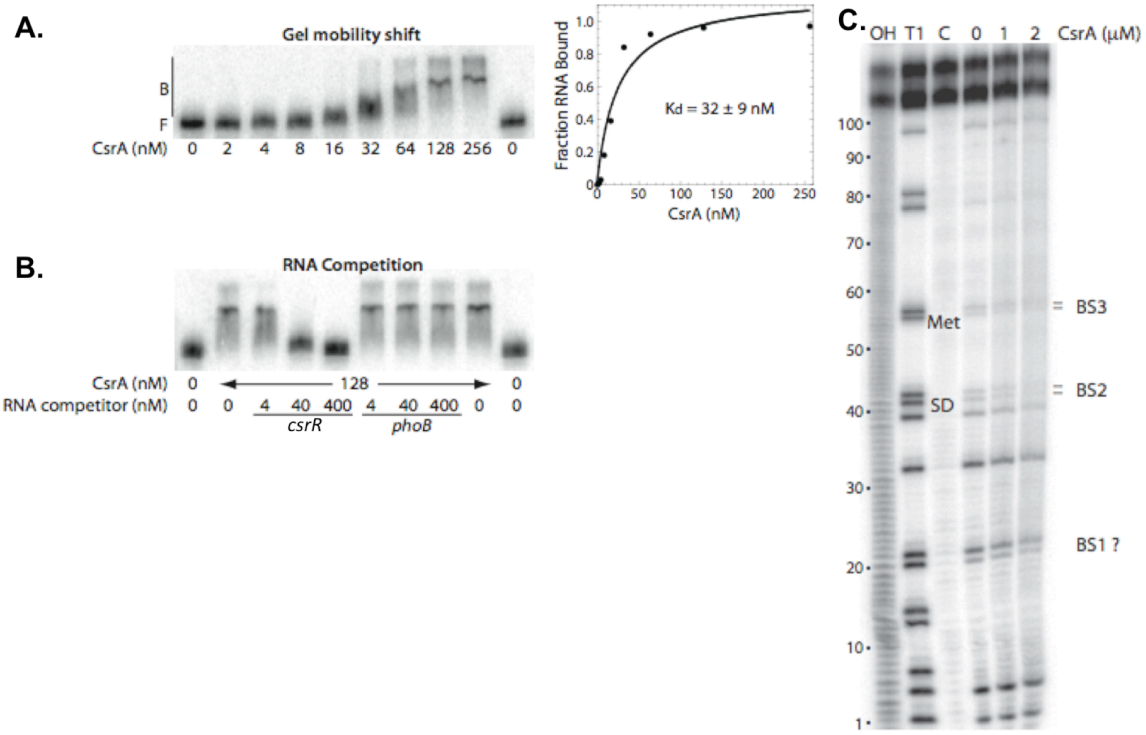

Figure 2.4. CsrA binds tightly and specifically to *csrR* RNA. (A) 5'-end-labeled *csrR* RNA (0.1 nM) was incubated with CsrA protein at the indicated concentration (nM) and the complexes separated by gel electrophoresis. Positions of bound (B) and free (F) RNA species are shown. The binding constant (K_d) was calculated from the binding curve shown at right. (B) Labeled *csrR* RNA (0.1 nM) was incubated with CsrA protein in the absence or presence of unlabeled specific (*csrR*) or non-specific (*E. coli phoB*) competitor RNA at the concentrations indicated. (C) 5'-end-labeled *csrR* RNA was treated with RNase T_1 in the presence of CsrA protein at the concentration indicated. Also shown are partial alkaline hydrolysis (OH) and RNase T_1 digestion (T1) ladders, as well as a control sample that lacked RNase T_1 treatment (C). Labels indicate residues for which CsrA protein reduced RNase T_1 cleavage (BS1-BS3), the *csrR* Shine-Dalgarno sequence (SD) and the translation initiation codon (Met). Numbering is with respect to the start site of *csrR* transcription (+1).

Figure 2.5

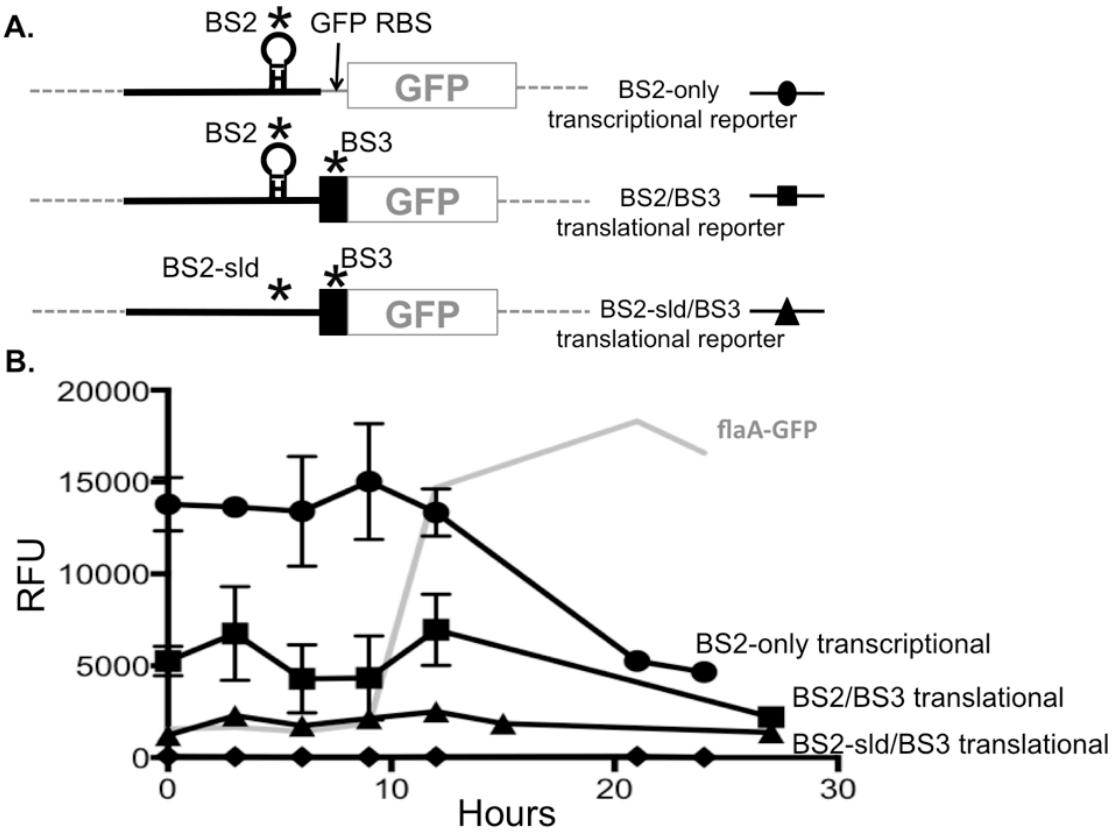

Figure 2.5. BS3 mediates repression of *csrR*-GFP reporters. (A) Schematics of *csrR*-GFP reporters to analyze *csrR* regulation. The BS2-only transcriptional reporter (MB1375, circles) encodes ~300 bp 5' to the translational start sequence, including the *csrR* SD, which overlaps the CsrA binding site BS2, and the *gfp* ribosome binding site (RBS). In the BS2/BS3 translational reporter (MB1376, squares) the *gfp* RBS is replaced by the first five codons of *csrR*, which encode the BS3 CsrA binding site. The BS2-sld/BS3 translational reporter (MB1377, triangles) was constructed by eliminating the stem-loop base-pairing to disrupt the BS2 CsrA binding site on the dual reporter. (B) E phase cultures of *L. pneumophila* carrying the reporter indicated were incubated for 24-28 h, and GFP fluorescence was quantified at the times indicated and expressed in relative fluorescence units (RFU). Reference strains were *L. pneumophila* carrying the empty *gfp* vector (diamonds) or *flaA*-GFP reporter (grey line) as an indicator of the transition into PE phase. Shown are means calculated from triplicate samples +/- SD; error bars for the stem-loop mutant are too small to be detected. A similar pattern was observed in at least two other experiments.

Figure 2.6

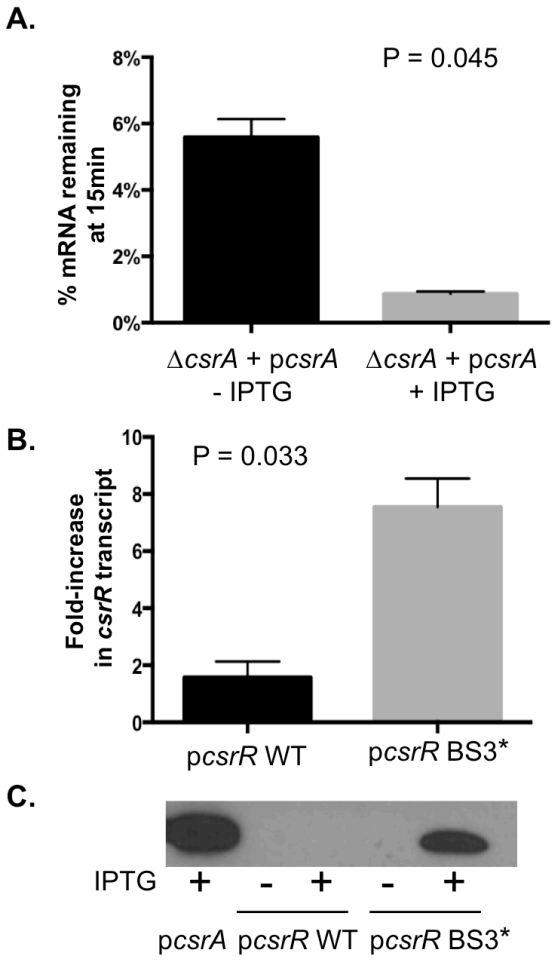

Figure 2.6. CsrA regulates *csrR* expression *in vivo.* (A) To analyze the impact of CsrA protein on *csrR* RNA stability, a conditional *csrA* mutant (MB464) was cultured without (black) or with (grey) IPTG to E phase, rifampin was added to inhibit new transcription, and RNA isolated 0 and 15 min later. Error bars indicate SD of duplicate samples, and significance was calculated using a student's *t*-test. Results are representative of two independent experiments. (B) *csrR* transcript stability was assessed via qRT-PCR of E phase *L. pneumophila* cultured without or with IPTG to induce expression of either WT (MB1378, black) or BS3 mutant *csrR* (MB1379, grey). Shown is the fold-increase in transcript in the induced strains relative to their uninduced controls. Error bars indicate SD from duplicate samples, and significance was calculated using a student's *t*-test. Results are representative of three independent experiments. (C) Strains shown in panel B were cultured into E phase with or without IPTG and then protein levels were analyzed by western assay using anti-His antibody. *L. pneumophila* encoding His-tagged CsrA on a plasmid served as the positive control. Data are representative of four independent experiments done with two independent isolates of each strain.

Figure 2.7

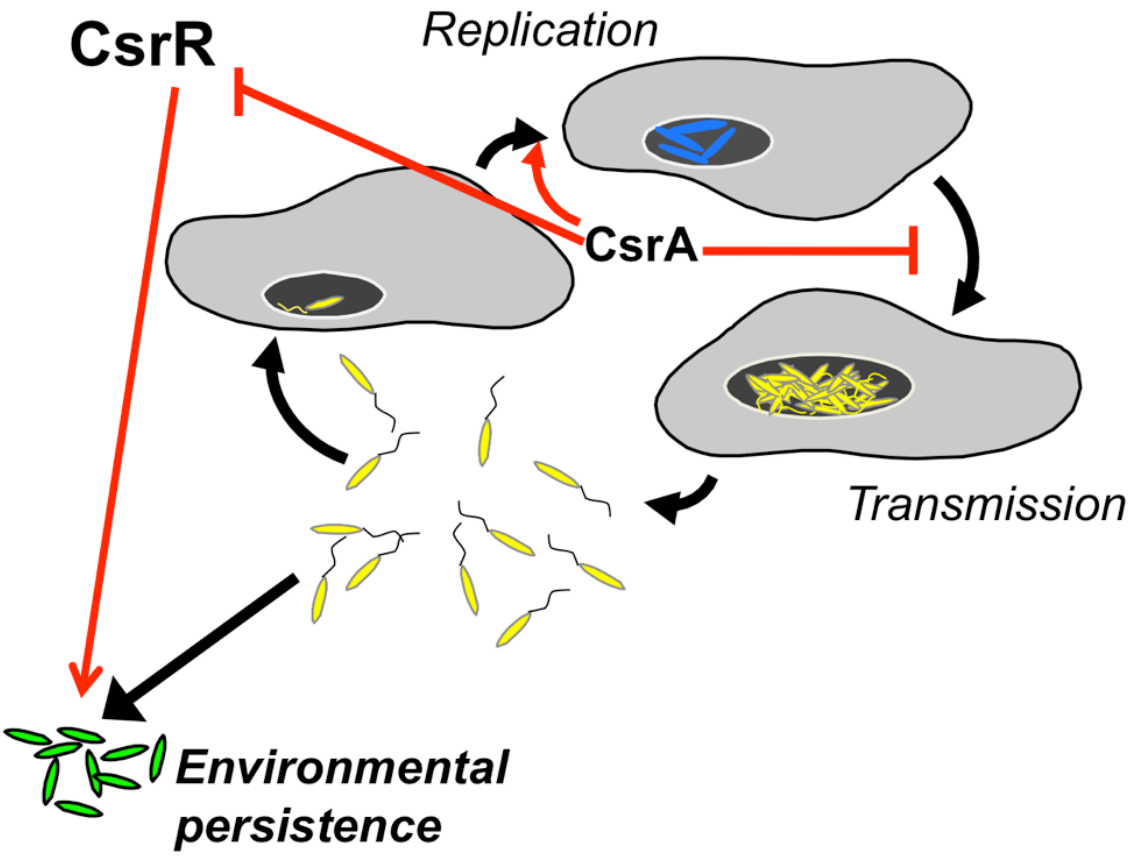

Figure 2.7. Model of CsrA and CsrR regulation of the *L. pneumophila* life cycle. Within host cells, CsrA induces transition into the replicative phase (blue bacteria) and represses both transmissive traits and *csrR* translation. When nutrients become limiting, CsrA repression is relieved, transcription of *csrR* declines, and *L. pneumophila* differentiates to the transmissive phase (yellow bacteria). After prolonged exposure to nutrient-poor water, CsrR promotes conversion to a resilient cell type (green bacteria).

CHAPTER III

CsrA paralog CsrT regulates core and accessory genome traits in *Legionella pneumophila.*

Abstract

Bacterial evolution is accelerated by mobile genetic elements. To spread horizontally and to benefit the recipient bacteria, genes encoded on these elements must be properly regulated. Among the legionellae are multiple Integrative Conjugative Elements (ICEs) that each encode a paralog of the broadly conserved regulator *csrA*. Using bioinformatic analyses, we deduced that specific *csrA* paralogs are co-inherited with particular lineages of the type IV secretion system that mediates horizontal spread of the ICE, suggesting a conserved regulatory interaction. To investigate the contribution of *csrA* to this class of mobile genetic elements, here we analyze the function of *csrT*, the *csrA* paralog encoded on *L. pneumophila* ICE-βox. CsrT abrogated the protection to hydrogen peroxide and macrophage degradation ICE-βox confers to *L. pneumophila*. In addition to modulating expression of some ICE-encoded cargo genes, CsrT also repressed *L. pneumophila* flagellin production and motility, a broadly conserved core genome trait that contributes to *L. pneumophila* transmission in water and virulence. Likewise, *csrT* repressed motility of *B. subtilis csrA* mutants, consistent with its predicted function as an mRNA-binding protein. That all known ICEs of legionellae

encode co-inherited *csrA* – type IV secretion system pairs suggests the capacity of CsrA superfamily proteins to regulate both ICE and host traits increases ICE promiscuity and expands *L. pneumophila* versatility.

Importance

ICEs are mobile blocks of DNA that encode the machinery for self-directed transmission among bacterial populations. ICEs typically carry their own regulatory genes to control expression of their transfer apparatus and cargo traits. In this study we show that all surveyed ICEs within the *Legionella* genus carry paralogs of the canonical life cycle regulator *csrA*, which can be classified into distinct families based on sequence. We then used the oxidative stress resistant ICE-βox and its encoded paralog *csrT* as a model to determine if there is a regulatory relationship between paralog and ICE. CsrT can negatively regulate not only accessory ICE traits but also the host bacterium's motility, a strategy that may increase the element's conjugal transmission in the environment.

Introduction

Mobile DNA elements can spread advantageous traits among bacterial populations and contribute to the evolution of pathogens (Wozniak et al. 2010). Integrative Conjugative Elements (ICEs) are one type of mobile transposable element that encode their own type IV secretion system (T4SS) for transfer between strains and also the enzymes necessary to integrate the element into the host chromosome. ICEs have expanded the genetic repertoire of a number of bacterial pathogens by carrying a variety of cargo genes that confer advantageous traits for

their host. Some confer resistance to antibiotics, metal ions and oxidative stress; others enhance nitrogen and chlorobenzoate metabolism, biofilm formation and host cell infection (Beaber et al. 2004; Rodriguez-Blanco et al. 2012; Flynn et al. 2014; Koraimann et al. 2014). Separate from the cargo genes are those encoding the elements' structural components, which are also diverse. Distinct classes of T4SSs mediate ICE transmission within and between species (Juhas et al. 2008). Such diversity may permit multiple ICEs to be maintained within one bacterial cell (Frost et al. 2010).

Bacteria coordinate expression of multiple extracellular structures, each with distinct purposes. *L. pneumophila* produces short pili for adherence or long T4SS conjugative pili for effector secretion or horizontal gene transfer, but not both pili at the same time (Stone et al. 1998). Likewise, simultaneous expression of similar T4SS by *L. pneumophila* is counter-productive, as activity of the core Dot/Icm T4SS is inhibited by the MobA component of a related RSF1010 conjugative plasmid (Segal et al. 1998). Accordingly, competition between ICEs likely provides selective pressure for regulatory circuits that ensure reciprocal expression of their T4SS machines.

The regulatory mechanisms governing the expression of ICE traits are as diverse as the ICEs themselves. For example, SXT from *Vibrio cholerae* and ICEBs1 from *Bacillus subtilis* encode regulatory genes whose products are cleaved or bound by host regulators induced by the bacterial SOS response (Beaber et al. 2004; Bose et al. 2008). Other ICEs are equipped with regulatory circuitries that act independently of host proteins. ICEclc from *Pseudomonas aeruginosa* and CTnDOT

from *Bacteriodes thetaiotaomicron* carry sensor proteins that respond to chlorobenzoate or tetracycline, respectively; these chemicals induce transcription of regulatory genes that drive expression of cargo loci that confer protection from these stressors (Sentchilo et al. 2003; Wang et al. 2004). Other ICEs, such as pLS20 from *B. subtilis* and pAD1 of *Enterococcus faecalis*, rely on secreted peptides or pheromones, respectively, to sense recipient cells in the environment and then induce a signaling cascade that activates transcription of the conjugative pilus (Bose et al. 2011; Clewell 2011). Finally, to regulate transfer, the *Agrobacterium tumefaciens* Ti conjugative plasmid responds to the secreted quorum sensing N-acyl-L-homoserine lactone (AHL) molecule (Haudecoeur et al. 2010).

The environmental and opportunistic pathogen *Legionella pneumophila* typically carries multiple ICEs in its genome. For instance, the Philadelphia-1 strain of *L. pneumophila* harbors three distinct ICEs (Rao et al. 2013). One of these mobile elements, ICE-βox, enhances the pathogen's resistance to bleach and the β-lactam antibiotics oxacillin and penicillin (Flynn et al. 2014). Strains that carry ICE-βox are also more infectious and tolerant of the macrophage phagocyte oxidase (Flynn et al. 2014). Protecting the bacterium from these stresses presumably benefits ICE-βox by ensuring a healthy host competent for spread of the element. Thus, expression of the ICE-βox-encoded T4SS transmission machinery and cargo traits must be regulated in a manner that ensures ICE-βox maintenance.

It is striking that ICE-βox encodes a paralog of *csrA,* a key regulator of *L. pneumophila* differentiation. The canonical CsrA is encoded in the "core" genome: it is ubquitous and highly conserved in the *Legionella* genus and is not encoded within

a horizontally acquired genetic element (Chien et al. 2004; Cazalet et al. 2008). CsrA is a broadly conserved mRNA binding protein that is essential for *L. pneumophila* replication and that represses numerous *L. pneumophila* transmissive phase traits, including expression of substrates of the core Dot/Icm T4SS (Molofsky et al. 2003; Nevo et al. 2014). Furthermore, similar to the dual CsrA system in *P. aeruginosa* (Marden et al. 2013), CsrA directly represses *csrR*, a newly identified *csrA* paralog in the *L. pneumophila* core genome, a design predicted to establish a reciprocal expression profile (Chapter II). The legionellae also encode a large "accessory" genome: highly variable regions of the genome frequently composed of horizontally acquired elements, including ICEs, that are not conserved between strains and species (Cazalet et al. 2004; Gomez-Valero et al. 2011; O'Connor et al. 2011). The ICE Trb-1 mobile element in the accessory genome of *L. pneumophila* strain Corby also encodes a *csrA* paralog, *lvrC* (Lautner et al. 2013). The *lvrC* locus encodes a repressor of ICE Trb-1 mobility, as deletion of the four-gene regulatory region significantly increased episomal ICE Trb-1 copies (Lautner et al. 2013). The phenotypic advantage of harboring ICE Trb-1 has not been identified; thus, whether *lvrC* also regulates ICE Trb-1 cargo genes remains to be determined.

Because paralogs of the known pluripotent regulator *csrA* are found in many *L. pneumophila* ICEs, and some ICEs – such as SXT, ICEBs1, and ICEclc discussed above – exhibit tight and reciprocal control, we sought to determine whether ICE *csrA* paralogs regulate their own core or cargo activities and competitive host traits. Given the observation that the *lvrC* locus represses ICE Trb-1 mobility and the canonical CsrA represses the core Dot/Icm T4SS, we postulated that *csrA* paralogs

function as repressors of ICE traits. Indeed, we demonstrate here co-inheritance of specific *csrA* paralogs with specific T4SS families encoded by ICEs. Using ectopic expression plasmids and the ICE-βox *csrA* paralog *csrT* as a model system, we show CsrT represses the oxidative stress resistance conferred by ICE-βox. In addition, CsrT expression in *L. pneumophila* or *B. subtilis* inhibited motility, a broadly conserved chromosomal pathway predicted to interfere with conjugation.

Materials and Methods

Bacterial strains, culture conditions and reagents. *Legionella pneumophila* strains were cultured in *N*-(2-acetamido)-2-aminoethanesulfonic acid (ACES; Sigma)-buffered yeast extract (AYE) broth supplemented with 100 μg/ml thymidine (Sigma) (AYET) at 37°C. Colony forming units (CFU) were quantified by plating on ACES-buffered charcoal yeast extract agar (CYE) supplemented with thymidine (CYET) and incubated at 37°C for four days (Feeley et al. 1979). Bacteria cultured overnight in AYET were diluted and grown overnight to obtain cells in exponential (E; OD_{600} of 1.2 to 1.8) or post exponential (PE; OD_{600} of 3.2 to 3.7) phases. When necessary for mutant or plasmid selection, the following antibiotics were added at the given concentrations: gentamicin (Gibco) at 10μg/mL; chloramphenicol (Fisher) at 5μg/mL. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from goldbio.com and added to a final concentration of 200μM unless otherwise indicated. Hydrogen peroxide (Sigma) was diluted in AYET and used at a concentration of 2mM. Strains, plasmids, and primers used in this study are summarized in Table 3.1.

Phylogeny and alignment. The predicted amino acid sequences of the 34 ICE-associated *csrA* paralogs (Table 3.2) were aligned using ClustalW2 (Larkin et al. 2007). The alignment was used to generate a maximum likelihood phylogenetic tree with phyML using the LG substitution model, 4 substitution rate categories, and 100 bootstraps (Guindon et al. 2010). The resulting tree was visualized using iTOL (Letunic et al. 2011).

Using the order of *csrA* paralogs determined from the phylogeny, pairwise alignment of the T4SSs were generated using tBLASTx and EasyFig (Sullivan et al. 2011). Nucleotide sequences of the T4SS and annotation were taken directly from the NCBI Nucleotide database (Table 3.3).

Induced expression constructs. Wild-type *csrA* and *csrT* with C-terminal 6x-histidine tags were amplified from Lp02 genomic DNA using primer sets ZA33+34 and ZA11+12, respectively. The resultant PCR products were digested with *Sal*I and *Hind*III and ligated into p206gent (Dalebroux et al. 2010). These plasmids were then electroportated into Lp02 to generate strains MB1389 and MB1383. We verified that adding 200μM IPTG resulted in protein expression by performing western blots using a 1:5000 dilution of anti-his(C-term)-HRP conjugated antibody following the western blot protocol described below (data not shown).

Mice. Six- to eight-week-old female A/J mice were purchased from Jackson Laboratories. Mice were housed in the University Laboratory Animal Medicine

Facility at the University of Michigan Medical School under specific-pathogen-free conditions. The University Committee on Use and Care of Animals approved all experiments conducted in this study.

Hydrogen peroxide exposure assay. Stress resistance was assessed by quantifying bacterial CFU after exposure to hydrogen peroxide. E phase strains containing pcsrT (MB1383, MB1384, MB1385) were cultured in AYET with or without IPTG induction overnight. After subculture to OD_{600} = 1.0, cultures were treated with 2mM H_2O_2 for 1 h at 37°C. Mean +/- SEM percent survival was calculated from triplicate samples.

Quantitative PCR. ICE-βox cargo gene expression was probed using quantitative Real-Time PCR (qRT-PCR). RNA from strain MB1383 cultured to E phase with or without 200μM IPTG was isolated using TRIZol reagent (Life Technologies) and its quality assessed using an Agilent Bioanalyzer 2100 according to manufacturer's protocol. Primer pairs KF162+163, KF152+153, and KF160+161 were used to quantify cargo gene expression in IPTG-induced RNA pools (*lpg2100, lpg2112 magA*, *lpg2098 msrA3*, respectively) relative to uninduced. As an internal control, chromosomal 16S rRNA levels were used to normalize input RNA levels using primer pair SL1+2. Log_2-transformed values of the mean ± standard error of the mean (SEM) fold change #(induced – uninduced) amplicons calculated from three independent experiments performed in triplicate are presented.

Intracellular growth and immunofluorescence microscopy. Growth of bacteria in bone marrow-derived macrophages from A/J mice (Jackson Laboratories) was assessed as described previously (Swanson et al. 1995). Strains were cultured to PE phase with or without 200μM IPTG overnight prior to infection. To account for infectivity differences, uninduced motile strains were infected at an MOI of 1 while induced nonmotile strains were infected at an MOI of 2. IPTG was included in the infection media of induced strains at a concentration of 1mM as described previously (Molofsky et al. 2003). Infection efficiency was quantified as the % CFU of bacteria associated with macrophages at 90 min post infection relative to input titer. Microscopy was performed by plating 3 x 10^5 macrophages on 12-mm glass coverslips overnight prior to infection with bacteria cultured as described above. At 90 min post infection, macrophages were fixed and stained as described previously (Swanson et al. 1996) using a 1:50 dilution of *L. pneumophila* primary antibody obtained from mouse monoclonal hybridoma cell line CRL-1765 (ATCC) and a 1:1000 dilution of anti-mouse IgG antibody conjugated to Oregon Green (Molecular Probes). The DNA stain DAPI (4′,6-diamidino-2-phenylindole) was purchased as ProLong Gold antifade reagent (Molecular Probes) and included in the mounting medium. NADPH oxidase experiments were performed using a pair of cell lines J774.16 (WT) and J774.D9 *phox* (deficient in gp91 subunit of NADPH oxidase) that were infected as described above (Goldberg et al. 1990).

Conjugation. To analyze ICE-βox transfer by conjugation, the thymidine auxotroph donor Lp02 strain containing a chloramphenicol-resistance marked ICE-βox and

pcsrT (MB1391) were cultured with or without IPTG overnight to E phase, and the thymidine prototroph recipient JR32 (MB370) was also grown overnight to E phase. 10^9 donor cells were mixed with 10^{10} recipient cells on 0.22-µm filters (Millipore) placed on prewarmed CYET agar plates and incubated for 2 h at 37°C as described previously (Vogel et al. 1998; Flynn et al. 2014). For induced strains, 10ul of 200mM IPTG was placed under the filter. Serial dilutions of the mating mixture were plated on CYET-cam (5µg/ml) to select for donors and CYE-cam to select for transconjugants. Transconjugation efficiency was calculated as CFU transconjugants/CFU donor cells in triplicate from three independent experiments.

Motility and flagellin western blot. Wild-type (MB110), Δ*flaA* mutant (MB1390), and strains carrying pcsrA and pcsrT (MB1389 and MB1383) were cultured in AYET to E-phase and then diluted to OD_{600} = 0.2. Strains MB1389 and MB1383 were split into two tubes, and one was induced with 200µM IPTG. All strains were then incubated at 37°C for ~15 h to PE phase, and then motility was assessed qualitatively by observation of wet mounts by inverted phase microscopy. Aliquots of the cultures were subsequently normalized to OD_{600} = 10, and then resuspended in 100µL Laemmli buffer (2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.005% bromophenol blue, 62.5mM Tris-HCL pH 6.8). Samples were then lysed by boiling for 5 min, and debris pelleted at 13,000 rpm for 2 min. Proteins were separated on a 12% mini-PROTEAN TGX precast gel (Bio-Rad), and Precision Plus Kaleidoscope Protein Standard ladder (Bio-Rad) was used as a size marker and to verify transfer. To verify equal loading and transfer of each sample, Ponceau S (Fisher) staining was

performed on a duplicate membrane. Flagellin was detected using a 1:100 dilution of 2A5 rabbit monoclonal antibody (Gift from NC Engelberg, UMMS), 1:3000 goat anti-rabbit secondary IgG-HRP (Pierce), and the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

***B. subtilis* growth conditions.** *Bacillus subtilis* strains were cultured in Luria-Bertani (LB) (10 g tryptone, 5 g yeast extract, 5 g NaCl per L) broth or on LB plates fortified with 1.5% Bacto agar at 37˚C. When appropriate, antibiotics were included at the following concentrations: 10 μg/ml tetracycline, 100 μg/ml spectinomycin, 5 μg/ml chloramphenicol, 5 μg/ml kanamycin, and 1 μg/ml erythromycin plus 25 μg/ml lincomycin. 1mM IPTG (Sigma) was added to the medium when appropriate.

***B. subtilis* strain construction.** To generate the *amyE::*$P_{hyspank}$*-csrALp* spec complementation construct, PCR product containing the *csrALp* coding region was amplified from *L. pneumophila* chromosomal DNA using the primer pair 3520/3521. DNA fragments containing the IPTG-inducible $P_{hyspank}$ promoter, a spectinomycin resistance cassette and the *amyE* region were amplified from the strain DS4940 as template using primers 3177/3519 and 3170/3180. Next, the PCR products were ligated using Gibson assembly (Gibson et al. 2009). The complementation constructs *amyE::*$P_{hyspank}$*-csrX* spec, *amyE::*$P_{hyspank}$*-csrT* spec, *amyE::*$P_{hyspank}$*-csrA-22* spec, and *amyE::*$P_{hyspank}$*-lvrC* spec were generated in a similar way using PCR amplicons from *L. pneumophila* chromosomal DNA using the primer pairs 3522/3523, 3524/3525, 3526/3527, and 3528/3529 respectively.

All constructs were first introduced into the domesticated strain DS2569 by natural competence and then transferred to the 3610 and DS6530 background using SPP1-mediated generalized phage transduction (Yasbin et al. 1974).

***B. subtilis* swarm expansion assay.** Cells were cultured to mid-log phase at 37˚C in LB and resuspended to OD_{600} = 10 in pH 8.0 PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4) containing 0.5% India ink (Higgins). Freshly prepared LB containing 0.7% Bacto agar (25 ml/plate) was dried for 20 min in a laminar flow hood, centrally inoculated with 10 µl of the cell suspension, dried for another 10 min, and incubated at 37˚C. The India ink demarks the origin of the colony, and the swarm radius was measured relative to the origin. For consistency, an axis was drawn on the back of the plate and swarm radii measurements were taken along this transect. For experiments including IPTG, cells were propagated in broth in the presence of IPTG, and IPTG was included in the swarm agar plates.

Results

***csrA* paralogs and their associated T4SSs are genetically linked.** The *Legionella* pan-genome has been predicted or demonstrated to encode many ICEs (Doleans-Jordheim et al. 2006; Wee et al. 2013; Flynn et al. 2014). In each of the 34 legionellae ICEs identified (Table 3.3), a *csrA*-like gene immediately precedes the locus encoding the T4SS. Previously we identified distinct families of *csrA*-like genes based on their amino acid sequence similarity (Chapter II). Likewise, *Legionella* genomic islands encode several different lineages of T4SSs (Segal et al. 1999;

Glockner et al. 2008; Lautner et al. 2013; Wee et al. 2013; Flynn et al. 2014). Therefore, we investigated whether particular lineages of ICE-encoded T4SS were genetically linked to a distinct *csrA* paralog. To do so, we first determined the relationship of all 34 *csrA* paralogs by generating a phylogenetic tree. Four distinct groups were apparent based on their amino acid sequence; only five of the 34 paralogs did not clearly fall into any group (Fig. 3.1A).

To determine if the T4SS lineages paired with a specific group of *csrA*-paralogs, we next aligned the T4SSs associated with each of the *csrA* paralogs. Indeed, within each *csrA*-paralog group, the associated T4SSs lineages were quite similar, whereas there was little similarity between T4SSs that are genetically linked to different *csrA* families (Fig. 3.1B). For instance, all the T4SSs associated with *csrA*-paralogs in group IV aligned strongly and indeed included every T4SS in the LGI lineage identified by Wee *et al.* (Wee et al. 2013). On the other hand, the T4SS in *csrA* group IV did not align with those of the Lvh T4SS lineage (Segal et al. 1999), the Trb T4SS lineage (Glockner et al. 2008), or the Tra T4SS lineage (Flynn et al. 2014). However, although the *csrA* paralogs were genetically linked to their T4SSs, the same relationship did not exist between *csrA* paralogs and associated cargo regions (data not shown). The linkage of particular groups of *csrA*-paralogs with specific T4SS lineages suggests selective pressure minimizes genetic drift between the *csrA* paralog and its associated T4SS locus. Accordingly, we postulated that there is a functional relationship between each ICE regulator and the genes that encode its conjugation machinery.

***csrT* inhibits ICE-βox transfer.** The genetic linkage between ICE-encoded *csrA* alleles and their adjacent T4SS loci appears to be maintained by selective pressure, and ICE conjugation is mediated by its T4SS (Lautner et al. 2013; Ilangovan et al. 2015). Therefore, we hypothesized that ICE transfer is regulated by ICE *csrA* genes, paralogs of the pluripotent *L. pneumophila* regulator CsrA. ICE-βox of *L. pneumophila* Philadelphia-1 carries the *csrA* paralog *lpg2094*, which we name here *csrT* (csrA paralog for ICE transfer) (Flynn et al. 2014). To test if ectopic expression of *csrT* is sufficient to regulate ICE-βox conjugative transfer, we created a plasmid (pcsrT) on which *csrT* transcription is induced by IPTG. To create a donor strain, and we electroporated pcsrT into Lp02, a lab derivative of the Philadelphia-1 strain that contains ICE-βox that has been marked with a chloramphenicol resistance cassette (Flynn et al. 2014). Donor cells cultured with or without IPTG were mated with the recipient strain JR32, a lab derivative of Philadelphia-1 that does not contain ICE-βox (Flynn et al. 2014). Indeed, ectopic expression of *csrT* in laboratory conditions inhibited conjugative transfer of ICE-βox, as transconjugation efficiency was reduced ~10 fold ($p < 0.05$) compared to donors that were not induced to express *csrT* (Fig. 3.2A). *csrT* repression of ICE-βox conjugation is consistent with the hypothesis that the *csrA* paralogs regulate their T4SSs, as predicted by their genetic relationship (Fig. 3.1) and the repression of ICE-Trb1 excision by the *lvrC* locus (Lautner et al. 2013)

ICE-βox excision is not sensitive to ectopic *csrT* expression.

Conjugative transfer requires that ICEs first direct their own excision from the host chromosome (Waldor 2010). Because ectopic *csrT* expression reduces ICE-βox conjugation and the locus encoding another *csrA* paralog, *lvrC*, inhibits excision of its own ICE Trb-1 (Lautner et al. 2013), we tested if *csrT* expression also represses ICE-βox excision. To quantify excised and integrated copies of ICE-βox, quantitative PCR using specific primers was performed as described previously (Flynn et al. 2014) using template DNA isolated from donor cells cultured to E and PE growth phases with or without ITPG. ICE-βox excision was not affected by ectopic expression of *csrT* in either growth phase (Fig 3.2B). Therefore, in the conditions tested ectopic expression of *csrT* was not sufficient to repress ICE-βox excision.

csrT represses _L. pneumophila_ resistance to hydrogen peroxide conferred by ICE-βox.

We postulated that CsrT inhibits the ICE-βox T4SS, based on three independent observations. ICE-encoded *csrA* paralogs are genetically linked with their associated T4SSs (Fig. 3.1); ectopic expression of *csrT* reduces ICE-βox conjugation but not excision (Fig. 3.2); and canonical CsrA inhibits Dot/Icm T4SS-mediated phenotypes, including host vesicular trafficking (Nevo et al. 2014). Therefore, we tested whether CsrT can repress ICE-βox-mediated phenotypes. In particular, we assayed the capacity of ICE-βox to increase *L. pneumophila* resistance to hydrogen peroxide (Flynn et al. 2014) after transforming donor, transconjugant and recipient strains with plasmid pcsrT. As predicted, ectopic *csrT* expression abrogated the protective effects of ICE-βox to 1 h exposure to 2mM hydrogen peroxide (Fig. 3.3A). Donor and

transconjugant ICE-positive strains that expressed *csrT* were ~10-fold more susceptible to killing by hydrogen peroxide when compared to the same strains cultured without IPTG-induction of *csrT* ($p < 0.05$, $p < 0.01$). *csrT* altered *L. pneumophila* sensitivity to oxidative stress by a mechanism that requires ICE-βox, since the strain that lacked the element was sensitive to hydrogen peroxide whether or not expression of pcsrT was induced with IPTG. Thus, when expressed ectopically, *csrT* inhibits ICE-βox-mediated hydrogen peroxide resistance.

ICE-βox cargo genes encoding oxidative-stress resistance enzymes are not repressed by ectopic *csrT* expression.

ICE-βox carries 38 putative cargo genes, some of which are predicted to confer resistance to oxidative stress (Flynn et al. 2014). In particular, three operons near the 3' end of the element are homologous to loci in other species that are activated during oxidative stress (Flynn et al. 2014). Accordingly, we tested whether *csrT* repression of these cargo genes accounts for its inhibition of ICE-βox-mediated oxidative stress resistance (Fig. 3.3A). To analyze transcript levels of homologs to cytochrome C oxidase (*lpg2100*), alkyl hydroperoxide reductase (*magA*), and methionine sulfoxide reductase (*msrA3*), we analyzed RNA isolated from E phase cultures of ICE-βox donor cells that either did or did not ectopically express *csrT*. In contrast to our hypothesis, in response to *csrT* induction, *lpg2100* was upregulated ~7 fold ($P < 0.001$), *magA* was induced ~2 fold ($p < 0.005$) and *msrA3* transcript levels were unaffected (Fig. 3.3B). Therefore, in culture conditions favorable for replication, ectopic expression of *csrT* enhances, rather than represses, transcription

of this set of ICE-βox encoded cargo genes. Accordingly, ectopically expressed *csrT* inhibits ICE-βox-mediated oxidative stress resistance (Fig. 3.3A) by a mechanism other than transcriptional repression of these three cargo genes.

csrT represses *L. pneumophila* evasion of degradative lysosomes.

Like hydrogen peroxide resistance, ICE-βox confers protection to *L. pneumophila* from degradative lysosomes in macrophage infection (Flynn et al. 2014). Therefore, as an independent measure of *csrT* repression of ICE-βox-mediated resistance to oxidative stress, we quantified bacterial survival in macrophages, where *L. pneumophila* encounters oxidative radicals (Klebanoff 2005). Donor, transconjugant, and recipient strains ectopically expressing *csrT* were co-cultured with bone marrow derived mouse macrophages in the presence or absence of IPTG. After a 90 min infection, bacterial resistance to degradation was quantified by fluorescence microscopy. As expected, bacterial cells from uninduced ICE-positive donor and transconjugant cultures resisted degradation much like PE phase WT Lp02 cells (Fig. 3.3C). However, when these same strains ectopically expressed *csrT*, the bacteria were more frequently degraded (~4-5 fold decrease in % intact bacteria, $p < 0.001$), as evident by particulate *L. pneumophila* antigen scattered throughout the macrophage, similar to E phase WT Lp02 cells (Fig. 3.3C). In contrast, ICE-negative bacteria were not more frequently degraded after *csrT* induction. Therefore, similar to its effect on resistance to hydrogen peroxide, ectopic expression of *csrT* is sufficient to inhibit ICE-βox-conferred resistance to

macrophage degradation, further supporting the model that CsrT is a negative regulator of ICE-βox-encoded traits.

To verify our microscopic assay of bacterial survival in macrophages (Fig. 3.3C), we quantified cell-associated *L. pneumophila*. Bone marrow-derived mouse macrophages were infected with donor, transconjugant and recipient strains carrying pcsrT, and infectivity was quantified as the fraction of input bacteria that were viable and cell-associated after a 90 min macrophage infection. Ectopic expression of *csrT* markedly reduced infection efficiency in ICE-βox positive donor and transconjugant strains (Fig. 3.3D, $p < 0.005$). In addition, induction of pcsrT also modestly decreased the already poor infectivity of the ICE-negative recipient strain (Fig. 3.3D, $p < 0.001$). This observation suggested that, in addition to inhibiting ICE-βox-encoded traits, CsrT might also modulate components of the *L. pneumophila* core genome.

Induction of *csrT* inhibits *L. pneumophila* growth in macrophages.

L. pneumophila containing ICE-βox replicate more efficiently in macrophages than those not containing the element (Flynn et al. 2014). Because ectopic expression of *csrT* represses ICE-βox-mediated oxidative stress resistance in broth and in macrophages (Fig. 3.3), we predicted *csrT* would also inhibit intracellular growth of *L. pneumophila*. To test the impact of *csrT* on intracellular replication, bone marrow-derived A/J mouse macrophages were infected with the donor strain cultured with or without IPTG to induce *csrT* expression. After 48 h, bacterial yield was quantified and expressed as fold-increase relative to the number of intracelluar

bacteria present at 2 h, to account for differences in infectivity (Fig. 3.3D). As expected, the yield of intracellular bacteria was significantly reduced (~100 fold, $p < 0.005$) when *csrT* was ectopically expressed (Fig. 3.4A).

In macrophages, ICE-βox confers *L. pneumophila* protection to oxidative stress generated by the NADPH oxidase (Flynn et al. 2014). To test whether the growth defect of strains induced to express *csrT* was due to the macrophage phagocyte oxidase, J774.16 WT and J774.D9 NADPH oxidase mutant macrophage cells were infected with the donor strain cultured with or without IPTG, and the yield of intracellular bacteria after 48 h was quantified. Much like the A/J mouse macrophage infection, the bacterial yield in J774.16 WT macrophages was significantly reduced by ectopic expression of *csrT* (~80 fold less, $p < 0.005$, Fig. 3.4B). Moreover, *csrT* also reduced the yield of *L. pneumophila* in the J774.D9 NADPH oxidase mutants (~50 fold less, $p < 0.005$, Fig. 3.4C). Therefore, ectopic expression of *csrT* inhibits *L. pneumophila* infection in macrophages by mechanism(s) that are not solely dependent on either the macrophage NADPH oxidase (Fig. 3.4C) or ICE-βox (Fig. 3.3D).

csrT* inhibits motility and flagellar assembly in *L. pneumophila. Efficient infection of macrophages by *L. pneumophila* requires not only evasion of the macrophage oxidative burst and phagosome-lysosome fusion (Berger et al. 1994; Swanson et al. 1996), but also motility, which increases bacterial contacts with host phagocytes (Molofsky et al. 2005). Our observation that ectopic expression of *csrT* decreased infectivity in a strain that lacks ICE-βox (Fig. 3.3D) and replication in

macrophages that lack NADPH oxidase (Fig. 3.4C) suggests an effect of CsrT on the core *L. pneumophila* genome. Since motility is a broadly conserved trait encoded by the core genome that enhances infectivity, we investigated whether *csrT* expression represses motility. We were further motivated by the fact that canonical CsrA represses not only Dot/Icm T4SS activity but also motility (Molofsky et al. 2003; Rasis et al. 2009; Nevo et al. 2014), indicating that *L. pneumophila* motility and T4SSs can belong to the same regulon. Indeed, when observed under a phase microscope, only ~ 5% of PE phase *L. pneumophila* that ectopically expressed *csrT*, were motile, as compared to approximately 90% of cells that were not induced to express *csrT* (Fig. 3.5A).

To examine *csrT* inhibition of motility in more detail, we analyzed flagellin levels by western analysis of *L. pneumophila* cultured with and without *csrT* induction. As a positive control for flagellin repression, we induced ectopic expression of canonical *csrA*, which represses motility in *L. pneumophila* (Segal et al. 1999; Molofsky et al. 2003). Similar to *csrA*, induction of *csrT* decreased flagellin protein levels (Fig. 3.5A), indicating that both CsrA and CsrT inhibit flagellin production.

***csrT* inhibits motility in *B. subtilis* by a conserved RNA-binding mechanism.** As in *L. pneumophila*, the *B. subtilis* CsrA protein inhibits bacterial motility (Mukherjee et al. 2011). Because the CsrA-binding site on its specific RNA target, *hag*, is defined in *B. subtilis*, we utilized this system to determine if CsrT can inhibit motility by the same molecular mechanism as the *B. subtilis* CsrA. In particular, we exploited a *B.*

subtilis mutant that is insensitive to CsrA repression of motility due to a point mutation in CsrA binding site on the *hag* mRNA (Mukherjee et al. 2011). Swarming motility by *B. subtilis* wild-type and Δ*fliW sow3* mutant strains that ectopically expressed *L. pneumophila csrA* or *csrT* was quantified. Induction of either *csrA* or *csrT* each inhibited motility in wild-type *B. subtilis*, verifying the assay is a sensitive readout of CsrA function. Neither *B. subtilis* CsrA nor *L. pneumophila* CsrT inhibited motility in the Δ*fliW sow3* mutant, which lacks the repressor's mRNA binding site. Since a single point mutation in the *hag* mRNA abrogated CsrT repression of motility, this CsrA paralog encoded by ICE-βox likely binds *hag* mRNA and represses its translation, consistent with a function as an RNA binding protein that recognizes the CsrA consensus sequence. In contrast, induction of canonical *L. pneumophila* CsrA inhibited *B. subtilis* swarming by a mechanism independent of the binding site defined for CsrA of *B. subtilis* (Fig. 3.5D, E). Therefore, the CsrA encoded by the core *L. pneumophila* genome inhibits motility in *B. subtilis* by a different mechanism than the *B. subtilis* CsrA,

Discussion

Legionellae must adapt to assault from diverse environmental and intracellular stresses. Horizontally-acquired ICEs may increase the fitness of this opportunistic pathogen (Miyamoto et al. 2003; Cazalet et al. 2004; Bandyopadhyay et al. 2007; Gomez-Valero et al. 2011; O'Connor et al. 2011; Flynn et al. 2014). Here we demonstrate that every *L. pneumophila* ICE identified to date encodes a *csrA*-like locus that is genetically linked to a particular T4SS lineage. Further, our genetic

analysis of *L. pneumophila* ICE-βox reveals that its *csrA* paralog *csrT* is a versatile regulator that cannot only repress the conjugative transfer and oxidative stress resistance encoded by the element, but also inhibit macrophage infection and bacterial motility when expressed ectopically.

By exploiting the well-characterized *B. subtilis* swarming motility pathway, we demonstrate that CsrT is likely an mRNA binding protein. Ectopic expression of *csrT* inhibited swarming motility in *B. subtilis* by a mechanism strictly dependent on the canonical CsrA binding site on the mRNA encoding flagellin (Fig. 3.5D). Based on the observation that a point mutation of the *B. subtilis* CsrA binding site abrogated repression by ICE-βox *csrT*, we deduce that CsrT protein binds mRNA and specifically recognizes the ANGGA consensus sequence for CsrA binding that is used by several bacterial species (Dubey et al. 2005; Mukherjee et al. 2011; Duss et al. 2014). Indeed, the CsrT protein is endowed with many of the key residues that equip *E. coli* CsrA to bind mRNA (Mercante et al. 2009)(Chapter II). Together our genetic and bioinformatic data strongly suggest that, like canonical CsrA, CsrT is a post-transcriptional regulator. Future biochemical experiments can verify that CsrT protein directly interacts with mRNA targets and identify CsrT binding sites in *L. pneumophila*.

Our bioinformatic data suggest that an RNA target of CsrT may lie in its associated T4SS locus. Sequence alignment of 34 ICE-associated *csrA*-paralogs revealed at least four distinct gene families (Fig. 3.1A). A combination of gene duplication events and subsequent genetic drift may have increased the size and diversity of this family of regulators. However, the high degree of similarity within

each of the four distinct *csrA* groups may indicate that ICEs depend on their paralog to regulate specific pathways. The fact that each *csrA*-paralog cluster specifically correlates with one of four previously identified T4SS lineages – Lvh (Segal et al. 1999), Trb (Glockner et al. 2008), Tra (Flynn et al. 2014), and LGI (Wee et al. 2013) (Fig. 3.1B) – predicts a functional relationship between ICE *csrA* paralogs and T4SSs. Whereas the specific amino acid sequence of each *csrA* paralog appears to be genetically constrained by the composition of the adjacent T4SS locus, the same cannot be said for the associated cargo regions, which are diverse. Therefore, we favor the model that *csrA* paralogs regulate the core machinery of their T4SSs, rather than ICE cargo genes.

The prediction that ICE *csrA* paralogs regulate phenotypes conferred by their T4SSs is consistent with the capacity of canonical CsrA to regulate myriad *L. pneumophila* traits, including *dot/icm* T4SS protein substrate secretion (Molofsky et al. 2003; Nevo et al. 2014). Indeed, we observed that ectopic expression of *csrT* reduced ICE-βox transfer ~ 10-fold (Fig. 3.2). Consistent with regulation of the T4SS machinery by CsrT, we confirmed that the reduced conjugation could not be accounted for by a defect in ICE-βox excision (data not shown), unlike the impact of the *csrA* locus encoded by the Trb-1 ICE (Lautner et al. 2013). In addition, although ectopic expression of *csrT* inhibited ICE-βox-mediated resistance to hydrogen peroxide (Fig. 3.3A), this repression was not due to reduced transcript levels of putative effectors in the cargo region (Fig. 3.3B). Therefore our phenotypic and bioinformatic data and knowledge of regulation by canonical CsrA (Segal et al. 1999) are consistent with the model that CsrT regulates its T4SS machinery.

In addition to repression of the ICE-βox-mediated phenotypes (Fig. 3.3A, C), *csrT* also inhibited macrophage infection in part by one or more mechanisms independent of ICE-βox. When expressed ectopically, *csrT* reduced the infectivity and intracellular yield of *L. pneumophila* whether or not the strain carried ICE-βox (Fig. 3.3D, data not shown). Furthermore, *csrT* reduced the bacterial yield in the J774.D9 phagocyte oxidase mutant cell line that is permissive for growth of *L. pneumophila* that do or do not encode ICE-βox (Flynn et al. 2014) (Fig. 3.4C). The observation that ectopic expression of *csrT* repressed macrophage infections independently of ICE-βox suggests that CsrT may regulate components of the core genome.

Indeed, CsrT inhibits flagellin expression and motility (Fig. 3.5A), a broadly conserved trait that contributes to *L. pneumophila* virulence (Molofsky et al. 2003). Reduced motility likely contributes to CsrT's inhibition of infectivity (Fig. 3.3D, 3.5A) – amotile *L. pneumophila* are less infectious in cell culture due to decreased contacts with macrophages (Molofsky et al. 2005). CsrT repression of motility may increase ICE-βox fitness because a stationary host conjugates more efficiently; alternatively, its inhibition of flagellin production may simply reflect its ancestry as a homologue of canonical CsrA, a repressor of *L. pneumophila* motility (Molofsky et al. 2003). On the other hand, ectopic expression of two other ICE-associated *csrA*-paralogs was not sufficient to inhibit *L. pneumophila* motility (*lpg1257* and *lpg1003*, data not shown). Accordingly, we speculate that genetic drift of individual ICE *csrA* paralogs may accommodate distinct regulatory demands of their corresponding T4SS class (Fig. 3.1), concomitantly relaxing their capacity to repress motility.

The observation that ectopic expression of *csrT* inhibits phenotypes as diverse as ICE-βox conjugation and host motility provides insight into why *csrA* regulators are ubiquitous among *L. pneumophila* ICEs. CsrA is a pluripotent post-transcriptional repressor of transmission traits that *L. pneumophila* requires for replication *in vitro* and in macrophages and amoebae (Molofsky et al. 2003). It follows that the capacity of a *csrA* paralog to manipulate a core life-cycle switch of its host bacterium would increase the fitness of these horizontally-acquired conjugal elements. A post-transcriptional regulator would equip the ICE to load its host cell with a pool of transcripts whose translation can be quickly de-repressed when environmental conditions favor the transfer of an ICE.

Contribution statement

Chapter III is the adaptation of a manuscript draft produced with significant contributions by both Dr. Kaitlin J. Flynn and myself. I proposed the project and designed most experiments. Figures 3.1 and 3.5A were performed by me. Figure 3.2 was a jointly performed by Dr. Flynn and me. Figure 3.3 and 3.4 were performed by Dr. Flynn and Brenda G. Byrne. Figure 3.5B was performed by Dr. Sampriti Mukherjee. Data analysis and writing was shared equally between Dr. Flynn and me.

Table 3.1. Strains, plasmids, and primers

Strains

Strain number	Name	Description	Source
E. coli			
MB1001	DH5α	F-endA1 *hsdR17* (r- m+) *supE44 thi-1 recA1 gyrA* (Nal^r) relA1 Δ(lacZYA-argF)$_{U169}$Φ80dLacZΔM15λpirRK6	Lab collection
MB1380	DH5α pcsrT		This study
MB1214	DH5α p206gent		(Dalebroux et al. 2010)
L.pneumophila			
MB110	Lp02	Philadelphia-1 Str^r HsdR− Thy−, wild-type strain for this study	(Berger et al. 1993)
MB1353	Lp02 ICE-βox::cam	MB110 with ICE-βox marked with camR cassette	(Flynn et al. 2014)
MB370	JR32	Philadelphia-1 Sm^r r^- m^+	(Wiater et al. 1994)
MB1355	JR32+ICE-βox::cam	MB370 with ICE-βox marked with camR cassette	(Flynn et al. 2014)
MB1383	Lp02 pcsrT	MB110 with IPTG-inducible *csrT*; referred to in paper as "donor" (except figure 2)	This study
MB1391	Lp02 ICE-βox::cam pcsrT	MB1353 with IPTG-inducible *csrT*; "donor" in figure 2	This study
MB1384	JR32 pcsrT	MB370 with IPTG-inducible *csrT*; referred to in paper as "recipient"	This study
MB1385	JR32+ICE-βox::cam pcsrT	MB1355 with IPTG-inducible *csrT*; referred to in paper as "transconjugant"	This study
MB1389	Lp02 pcsrA	MB110 with IPTG-inducible *csrA*	This study
MB1390	Lp02 Δ*flaA*	Lp02 *flaA* deletion mutant	Lab collection
B. subtilis			
3610	Wild type		
DS6530	Δ*fliW sow3*	3610 with *fliW* deletion and a point mutation in the *hag* leader sequence, making it insensitive to CsrA repression of motility	(Mukherjee et al. 2011)
DS4940	*amyE::*$P_{hyspank}$*-csrABs spec*	3610 with IPTG-inducible *B. subtilis csrA*	(Mukherjee et al. 2011)
DK675	*amyE::*$P_{hyspank}$*-csrALp spec*	3610 with IPTG-inducible *L. pneumophila csrA*	This study
DK676	*amyE::*$P_{hyspank}$*-csrR spec*	3610 with IPTG-inducible *L. pneumophila csrR*	This study
DK677	*amyE::*$P_{hyspank}$*-csrA-T spec*	3610 with IPTG-inducible *L. pneumophila csrT*	This study
DK678	*amyE::*$P_{hyspank}$*-csrA-22 spec*	3610 with IPTG-inducible *L. pneumophila csrA-22 (lpg1003)*	This study
DK679	*amyE::*$P_{hyspank}$*-lvrC spec*	3610 with IPTG-inducible *L. pneumophila LvrC (lpg1257)*	This study
DK1469	*ΔfliWcsrA amyE::*$P_{hyspank}$*-csrABs spec*	DS6530 mutant with IPTG-inducible *B. subtilis csrA*	This study
DK1470	*ΔfliWcsrA amyE::*$P_{hyspank}$*-csrALp spec*	DS6530 with IPTG-inducible *L. pneumophila csrA*	This study
DK1471	*ΔfliWcsrA amyE::*$P_{hyspank}$*-csrT spec*	DS6530 with IPTG-inducible *L. pneumophila csrT*	This study
DK1472	*ΔfliWcsrA amyE::*$P_{hyspank}$*-lvrC spec*	DS6530 with IPTG-inducible *L. pneumophila lvrC (lpg1257)*	This study

Plasmids

Plasmid	Description	Source
p206gent	pMMB66EH derivative, Δ*mob*, *lacI*^{q}, $P_{taclacUV5}$, $gent^R$	(Dalebroux et al. 2010)
pcsrA	p206gent with *csrA* with 6x-his epitope ligated into the SalI and HindIII sites, IPTG-inducible *cs*	Chapter II
pcsrT	p206gent with *csrT* with 6x-his epitope ligated into the SalI and HindIII sites, IPTG-inducible *cs*	This study

Primers			
Number	description	sequence	Source
ZA11	Fwd SalI *csrT*	GTCGACCTGAAGTGGCTAAAACCCAATTAC	This study
ZA12	Rev HindIII-6x his-*csrT*	AAGCTTTTAATGATGATGATGATGATGATGGTGATTGAGGATTGATTGCT	This study
ZA33	SalI *csrA*-his fwd	GTCGACGGAAGCGTTCTATGGAGAAGTAGG	Chapter II
ZA34	HindIII *csrA*-his rev	AAGCTTTTAATGATGATGATGATGATGTACTGCTTGTTCCGAATCATCAG	Chapter II
KF152	*lpg2112* fwd	CACTGGAAGCCTACATTGGG	This study
KF153	*lpg2112*rev	CCATTATAACCTGCTGAGACC	This study
KF160	*lpg2098* fwd	GGTTTTGTACTGTTGGGTCG	This study
KF161	*lpg2098* rev	CTATGATGGTGGTACTGAGCC	This study
KF162	*lpg2100* fwd	GTGATACCCGCACTGTGG	This study
KF163	*lpg2100* rev	CAAACACAACACCCACCTC	This study
SL1	16S fwd	AGAAGGCCTGAGGGTTGTAAAGCA	lab collection
SL2	16S rev	ACCCTTTACGCCCAGTAATTCCGA	lab collection
3170	*amyE* fwd	GGTGGAAACGAGGTCATCATTTC	This study
3177	Spectinomycin fwd	CTAATTCAAGGCGTGTCTCAC	(Mukherjee et al. 2015)
3180	*amyE* rev	GCGGTATTCCGTATGTCAAG	(Mukherjee et al. 2015)
3519	Spectinomycin rev	CACTCTACCTCCTGCTAGCT	This study
3520	*csrALp* fwd	AGCTAGCAGGAGGTAGAGTGATGTTGATTTTGACTCGGCGT	This study
3521	*csrALp* rev	ATGATGACCTCGTTTCCACCAGTTCCCACAGCTTATATTAATT	This study
3524	*csrT* fwd	AGCTAGCAGGAGGTAGAGTGATGTTGAGTTTAACTAGAAGAGT	This study
3525	*csrT* rev	ATGATGACCTCGTTTCCACCTTAGAAAAATAGCCATCGCCC	This study
3526	*csrA-22* fwd	AGCTAGCAGGAGGTAGAGTGATGCTGATCCTAGACCGCAA	This study
3527	*csrA-22* rev	ATGATGACCTCGTTTCCACCCGTCGGCTATAGTTAAAGGAC	This study
3528	*lvrC* fwd	AGCTAGCAGGAGGTAGAGTGATGTTGGTTTTAACACGAAAAGC	This study
3529	*lvrC* rev	ATGATGACCTCGTTTCCACCGGGCTTATTTCCAAATCGCTC	This study

Table 3.2. *csrA* paralogs in *Legionellae* ICEs

Gene	Species	Strain	Amino acid sequence
lpep00046	*L. pneumophila*	LPE509	MLILTRRIGETVVIGDDVFITILGIKGNQIRLGFDAPDHVSIHRQEIYL KVQEQKKMRLDSGEAVNGNGMLITPLKQTEPHQYTAH
lpc0169	*L. pneumophila*	Corby	MLVLTRRVGESVVIHDDVYCTIVGYRDGEVRLAFDAPRSIPVHRDEI QRRIHRARIKDNWFIDKAANKESIVDRLINKFKNSTSAVKA
lpa00223	*L. pneumophila*	Alcoy	MLVLTRRVGESVVIHDDVYCTIVGYRDGEVRLAFDAPRSIPVHRDEI QRRIHRARIKDNWFIDKAANKESIVDRLINKFKNSTSAVKA
lpc2813	*L. pneumophila*	Corby	MLVLTRRVGESVVISEDIYCTIVGYQNDEVRLAFDAPKSIPIHRDEIQ RRIYRDQIKDNKFVDKAANNESIVDRLINKFKSSASPTNS
lpa00796	*L. pneumophila*	Alcoy	MLVLTRRVGESVVISEDIYCTIVGYQNDEVRLAFDAPKSIPIHRDEIQ RRIYRDQIKDNKFVDKAANNESIVDRLINKFKSSASPTNS
lpo2456	*L. pneumophila*	Lorraine	MLVLTRKVGESVVISEEVYCTVVGYRDGEVRLAFDAPQSIPVHRDEI QRRIYRERQRDQWFNDSPSNKENIVDRLISKFKNGLKSA
lpe01074	*L. pneumophila*	LPE509	MLVLTRKVGESVVISEEVYCTVVGYRDGEVRLAFDAPQSIPVHRDEI QRRIYRERQKDQWFSDSPSNKESIIDRLISKFKHGLKSA
llo2874	*L. longbeachae*	NSW150	MLVLTRKIGESVVISEDIYCTVVGYRDGEVRLGFDAPQSIPIHRDEIQR RIYRERQKDQSFNDSPPHKKSIVDRLISKFKHELKSA
plpl0012	*L. pneumophila*	Lens	MLILERRVGETVAIENKVFCMVLDHQLDGQLKLAFDAPECVPFHRF EIQEGS
plpp0016	*L. pneumophila*	Paris	MLILTRRVGDTVVIGNEVFCTVLEQQHDGQIKLAFDAPKSIPIHRFEI QKQIMQKIIEGTYTNDVAWNETVIERLTSQFNRVSHWN
lpop0154	*L. pneumophila*	Lorraine	MLILTRRIGETVLINDDIYITVLGVKGNQVRLGFDAPQDVIIHRQEIH QKIKKEQSLFLNKPGQWKEARGVQTH
lpg2094	*L. pneumophila*	Philadelphia-1	MLSLTRRVGESIVIGEDIFITVLCCKGNQVRIGFNAPNSVAIHRYEIYQ KIQSEKHDGLADPTKKFCPSMQQSILNHH
lpe01011	*L. pneumophila*	LPE509	MLSLTRRVGESIVIGEDIFITVLCCKGNQVRIGFNAPNSVAIHRYEIYQ KIQSEKHDGLADPTKKFCPSMQQSILNHH
lp12_2035	*L. pneumophila*	ATCC 43290	MLSLTRRVGESIVIGEDIFITVLCCKGNQVRIGFNAPNSVAIHRYEIYQ KIQSEKHDGLADPTKKFCPSMQQSILNHH
lpp1074	*L. pneumophila*	Paris	MLTLIRRIGEAIYIDKGRIKVHLISEKEGLIRLGIEAPKHVDIERKEVFV RKAVAQHEQAQALRNQSGGDDA
lp12_1033	*L. pneumophila*	ATCC 43290	MLTLIRRIGEAIYIDKGRIKVHLISEKEGLIRLGIEAPKHVDIERKEVFV RKAVAQHEQAQALRNQSGGDDA
lpo2781	*L. pneumophila*	Lorraine	MLILIRRMGEAIYIDKGRIKVLLISEKEGLIKLGIDAPKHIDVERKEVFI QKAMEQHALAQKLRDKSTESGGNHA
lpg1003	*L. pneumophila*	Philadelphia-1	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRAPSEIDIDRKEVFI RKYIQKLDQENKSNQG
lpl1036	*L. pneumophila*	Lens	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRASSEIDIDRKEVFIR KYIQKLDQENKSNQG
lpc2276	*L. pneumophila*	Corby	MLILDRKIGEEIFINKGKIKITVLYEKNGLIGIGVRAPSEIDIDRKEVFI RKYIQKLDQENKPNQE
lpa01531	*L. pneumophila*	Alcoy	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRASSEIDIDRKEVFIR KYIQKLDQENKSNQG
lpv1151	*L. pneumophila*	HL06041035	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRAPSEIDIDRKEVFI RKYIQKLDQENKSNQG
lpe02190	*L. pneumophila*	LPE509	MLILDRKIGEEIYINKGKIKITVLYEKNGLIGIGVRAPSEIDIDRKEVFI RKYIQKLDQENKSNQG
lpp2378	*L. pneumophila*	Paris	MLVLTRKKGEQIVIDKGQIEIHVIYQRRGVVALGIKAPAHIDVDRKEI FLRKQTNPQNNDKEISK
lpc1860	*L. pneumophila*	Corby	MLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKAPAHIDVDRKEI FLRKQTNPQNTNIEEPK
lpv1819	*L. pneumophila*	HL06041035	MLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKAPAHIDVDRKEI FLRKQTNPQNTDIEEPK
lp12_2077	*L. pneumophila*	ATCC 43290	MMLVLTRKKGEQILIDKGQIEIHVIYQRRGVVALGIKAPAHIDVDRKE IFLRKQTNPQNTNIEEPK
llb1171	*L. longbeachae*	D-4968	MLVLTRRVGEQIFIDKGQIQIKVLFVRNGNIALGIQAPPNVDVDREEI YYLKKEGIIVESN
lpg1257	*L. pneumophila*	Philadelphia-1	MLVLTRKAGQQILIGKGLIQMKVLKVDDDIISIGIKAPQHIDIDREEIY LKKLQQEQAESSMQKVAP
lpl0150	*L. pneumophila*	Lens	MLVLTRKAGQQILIGKGLIQMKVLKVDDDIISIGIKAPQHIDIDREEIY LKKLQQEQAESSMQKVAP
lpv0167	*L. pneumophila*	HL06041035	MLVLTRKAGQQILIGKGLIQMKVLKVDDDIISIGIKAPQHIDIDREEIY LKKLQQEQAESSMQKVAP
lpp0168	*L. pneumophila*	Paris	MLVLTRKAGQQILIGKGLIQMKVLKVENDIISIGIKAPSHIDIDREEIYF RKLRQEQEAANDAEMAI
llb1138	*L. longbeachae*	D-4968	MLILSRKIGENVLIDQGTIQIKLLDVKGRYARIGFIAPAGTDIDREEIYI RKKQSRDLNKKAANHEVNN

llo1850	*L. longbeachae*	NSW150	MLVLERKIGQKVVIDNGAIEVKVLKPHGDMIRLGFKAPQNMDINKE EIYLRKVLQVPEFLKPVVRNHETRRRK

Each gene is listed with its locus tag and inferred amino acid sequence.

Table 3.3. Summary of ICEs from Figure 3.1B.

csr gene marking ICE	Strain	Accession number	5' T4SS coordinate	3' T4SS coordinate	T4SS lineage (color in Fig. 1B)
lpc0169	*L. pneumophila* Corby	CP000675.2	184300	200451	Trb (blue)
lpa00223	*L. pneumophila* Alcoy	CP001828.1	184068	200219	Trb (blue)
lpc2813	*L. pneumophila* Corby	CP000675.2	617842	633964	Trb (blue)
lpa00796	*L. pneumophila* Alcoy	CP001828.1	613023	629145	Trb (blue)
lpo2456	*L. pneumophila* Lorraine	FQ958210.1	2581394	2597460	Trb (blue)
lpe01074	*L. pneumophila* LPE509	CP003885.1	1109051	1125125	Trb (blue)
llo2874	*L. longbeachae* NSW150	FN650140.1	3355412	3372833	Trb (blue)
lpg2094	*L. pneumophila* Philadeliphia-1	AE017354.1	2319208	2344884	Tra (red)
lpe01011	*L. pneumophila* LPE509	CP003885.1	1051761	1068953	Tra (red)
lp12_2035	*L. pneumophila* ATCC 43290	CP003192.1	2243991	2261276	Tra (red)
lpop0154	*L. pneumophila* Lorraine	FQ958212.1	122878	142676	Diverse (purple)
lpep00046	*L. pneumophila* LPE509	CP003886.1	19244	40410	Diverse (purple)
plpp0016	*L. pneumophila* Paris	CR628338.1	13345	42715	Diverse (purple)
plpl0012	*L. pneumophila* Lens	CR628339.1	10787	28932	Diverse (purple)
llo1850	*L. longbeachae* NSW150	FN650140.1	2164094	2181119	Diverse (purple)
llb1138	*L. longbeachae* D-4968	ACZG01000001.1	1196029	1211847	Lvh (yellow)
lpp0168	*L. pneumophila* Paris	CR628336.1	194054	210931	Lvh (yellow)
lpg1257	*L. pneumophila* Philadelphia-1	AE017354.1	1368713	1386591	Lvh (yellow)
lpl0150	*L. pneumophila* Lens	CR628337.1	181867	201513	Lvh (yellow)
lpv0167	*L. pneumophila* HL06041035	FQ958211.1	181295	198468	Lvh (yellow)
llb1171	*L. longbeachae* D-4968	ACZG01000001.1	1236013	1254272	Lgi (green)
lpc1860	*L. pneumophila* Corby	CP000675.2	2808352	2826507	Lgi (green)
lp12_2077	*L. pneumophila* ATCC 43290	CP003192.1	2296777	2314983	Lgi (green)
lpp2378	*L. pneumophila* Paris	CR628336.1	2726191	2744349	Lgi (green)
lpv1819	*L. pneumophila*	FQ958211.1	1840909	1859070	Lgi (green)

	HL06041035				
	L. pneumophila				
lpo2781	Lorraine	FQ958210.1	2899570	2917401	Lgi (green)
	L. pneumophila				
lp12_1033	ATCC 43290	CP003192.1	1101878	1119680	Lgi (green)
	L. pneumophila				
lpp1074	Paris	CR628336.1	1177462	1195264	Lgi (green)
	L. pneumophila				
lpv1151	HL06041035	FQ958211.1	1140510	1158271	Lgi (green)
	L. pneumophila				
lpg1003	Philadelphia-1	AE017354.1	1076121	1093906	Lgi (green)
	L. pneumophila				
lpe02190	LPE509	CP003885.1	2334332	2352138	Lgi (green)
	L. pneumophila				
lpc2276	Corby	CP000675.2	1189397	1207218	Lgi (green)
	L. pneumophila				
lpl1036	Lens	CR628337.1	1149440	1167201	Lgi (green)
	L. pneumophila				
lpa01531	Alcoy	CP001828.1	1185994	1203755	Lgi (green)

Sequence identification for the T4SSs aligned in Figure 1B, listed in the same order as they are aligned.

Figure 3.1

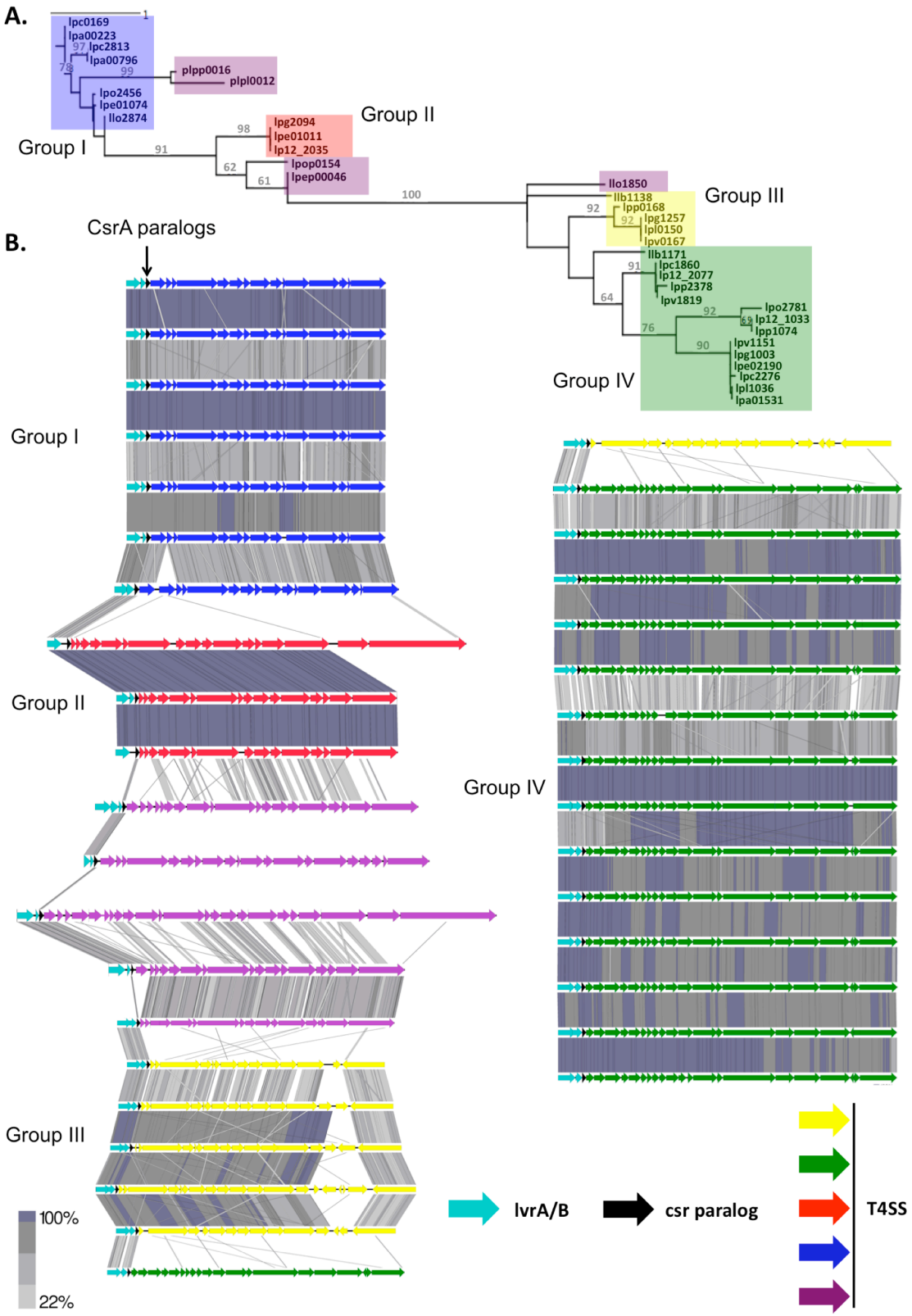
A.
lpc0169
lpa00223
lpc2813
lpa00796
lpo2456
lpe01074
llo2874
lpp0016
lpl0012
lpg2094
lpe01011
lp12_2035
lpop0154
lpep00046
llo1850
llb1138
lpp0168
lpg1257
lpl0150
lpv0167
llb1171
lpc1860
lp12_2077
lpp2378
lpv1819
lpo2781
lp12_1033
lpp1074
lpv1151
lpg1003
lpe02190
lpc2276
lpl1036
lpa01531
Group I
Group II
Group III
Group IV
B.
CsrA paralogs
Group I
Group II
Group III
Group IV
100%
22%
lvrA/B
csr paralog
T4SS

Figure 3.1. Relationship of CsrA paralogs and T4SSs. (A) To determine the relationship between 34 ICE-associated *csrA* loci, a maximum likelihood phylogenetic tree was generated from their predicted amino acid sequences. Four distinct clusters are boxed in blue, red, yellow, and green. The five protein sequences that did not clearly fall into any cluster are boxed in purple. Branches with at least 60% support from 100 bootstraps are labeled. (B) Pairwise translated nucleotide sequence identity (tBLASTx) of 34 T4SSs in the legionellae pan-genome was performed based on the relationship of their adjacent *csrA* paralogs as determined by the phylogeny in (A). *lvrA/B* genes are teal arrows; *csrA* paralogs are black arrows and indicated by pointer; T4SSs are colored and labeled by group to correspond with the boxes in (A). Grey bars between T4SS operons indicate the % amino acid identity as shown by the scale bar. The bottom two T4SSs from the left column are reproduced as the top two in the right column for continuity.

Figure 3.2

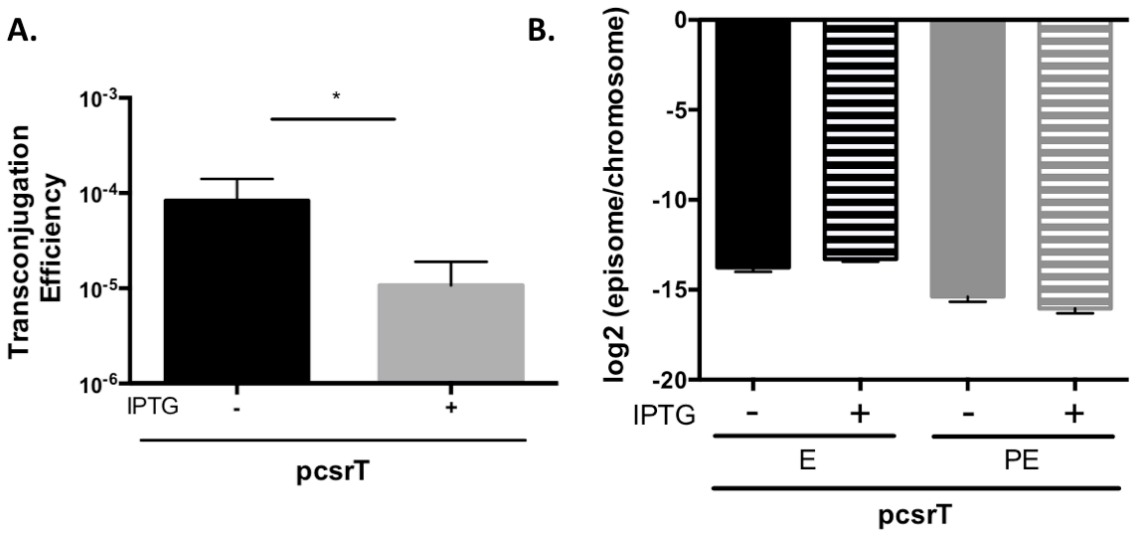

Figure 3.2: Ectopic *csrT* expression reduces ICE-βox transfer. (A) ICE-βox null recipient strain (MB370) was mated with E phase ICE-βox donor strain containing pcsrT (MB1383) after overnight culture without (black) or with IPTG (grey) to induce ectopic expression of CsrT. Mean transconjugation efficiency of ICE-βox is expressed as (CFU transconjugants/ CFU donor) +/- SEM of results from three independent experiments. *t*-test indicates the difference between uninduced and induced transconjugation efficiency is statistically significant ($p < 0.05$). (B) Strain MB1383 was grown into E phase (black) or PE phase (grey) without (solid) or with (striped) IPTG and genomic DNA was assessed by quantitative PCR for integrated (chromosomal) vs. excised (episomal) ICE-βox using primer pairs specific for each species. Data are expressed as the $\log_2$ (episomal form/chromosomal form) and represent the mean of three replicate qPCR reactions, and error bars indicate SD.

Figure 3.3

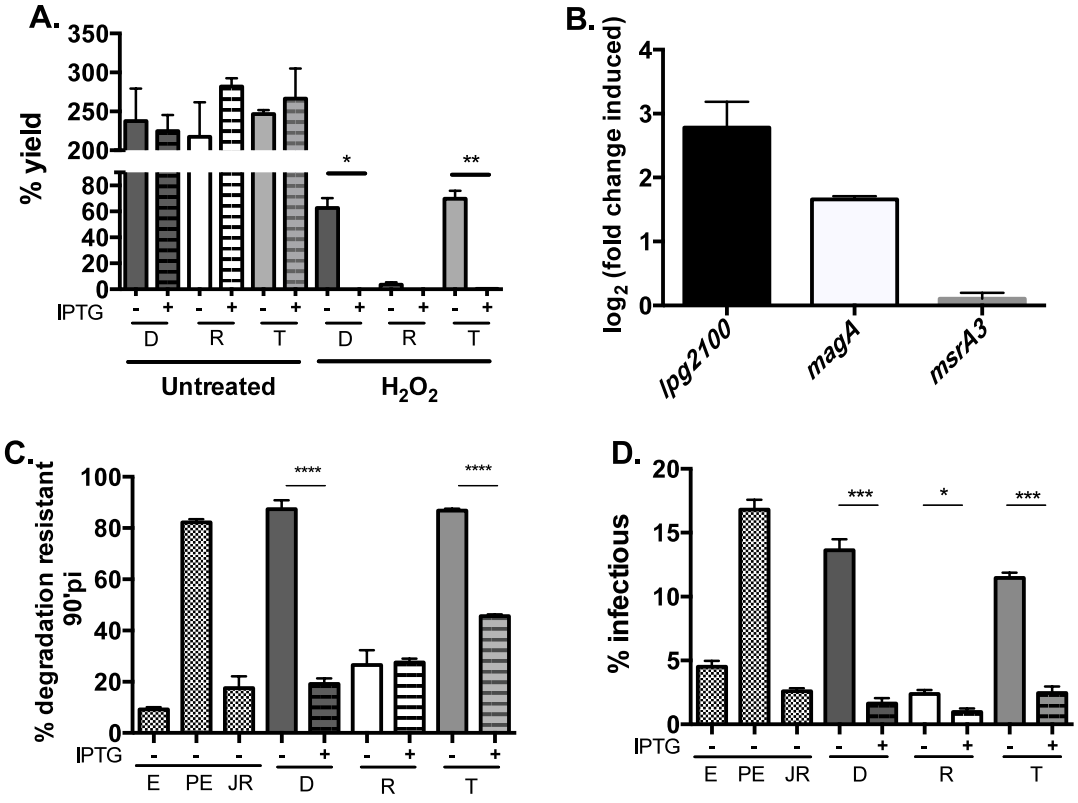

Figure 3.3: Ectopic *csrT* expression represses oxidative stress resistance conferred by ICE-βox. (A) CsrT abrogates the capacity of ICE-βox to protect *L. pneumophila* from hydrogen peroxide stress. E phase cultures of ICE-βox donor (D, dark grey), naïve recipient (R, white) or ICE-βox transconjugant (T, light gray) strains that contained IPTG-inducible plasmid pcsrT (Strains MB1383, MB1384, and MB1385, respectively) were cultured overnight with (striped) or without (plain) IPTG and then exposed to 2mM H_2O_2 for 1 h. Mean % survival +/- SEM was calculated from three independent experiments as (CFU post-treatment/CFU pre-treatment)*100. *t*-tests indicate statistically significant differences between treated, pcsrT-induced ICE-βox strains and the treated uninduced strains (*, $p < 0.05$, ** $p < 0.01$). (B) ICE-βox cargo transcripts are differentially expressed in *csrT*-induced strains. Expression analysis of three ICE-βox cargo genes – *lpg2100*, *magA*, and *msrA3* – was performed on RNA isolated from E phase ICE-βox donor cells carrying pcsrT (MB 1383) after culture overnight with or without IPTG. qRT-PCR results are expressed as log_2(transcript induced – transcript uninduced) and means +/- SEM of three independent RNA isolates are presented. (C) CsrT induction represses bacterial degradation resistance in macrophages. E and PE phase strain MB110 Lp02 and PE phase strain MB370 JR32 (E, PE, and JR, respectively; hatched bars), ICE-βox donor (D, dark grey), recipient (R, white), or transconjugant (T, light gray) strains containing the inducible pcsrT plasmid (Strains MB1383, MB1384, and

MB1385)) that had been grown overnight with (striped) or without (plain) IPTG, were used to infect A/J macrophages at an MOI of 2 (induced) or 1 (uninduced). At 90 min post infection, bacterial integrity was quantified by immunofluorescence microscopy. Presented are the mean % +/- SEM of total bacteria that are intact vs. degraded from three independent experiments. *t*-tests indicate that differences between induced and uninduced D and T strains are significant ($p < 0.001$). (D) CsrT induction inhibits efficient infection of macrophages. A/J mouse macrophages were infected as described in (C). At 90 min post infection, cells were lysed and CFU enumerated. Percent of infectious bacteria was calculated as (CFU at 90'/CFU of infection input) * 100 and is expressed as mean +/- SEM from three replicates in one experiment representative of three others. *t*-test confirm differences between induced and uninduced strains are significant (* $p < 0.005$, *** $p < 0.05$).

Figure 3.4

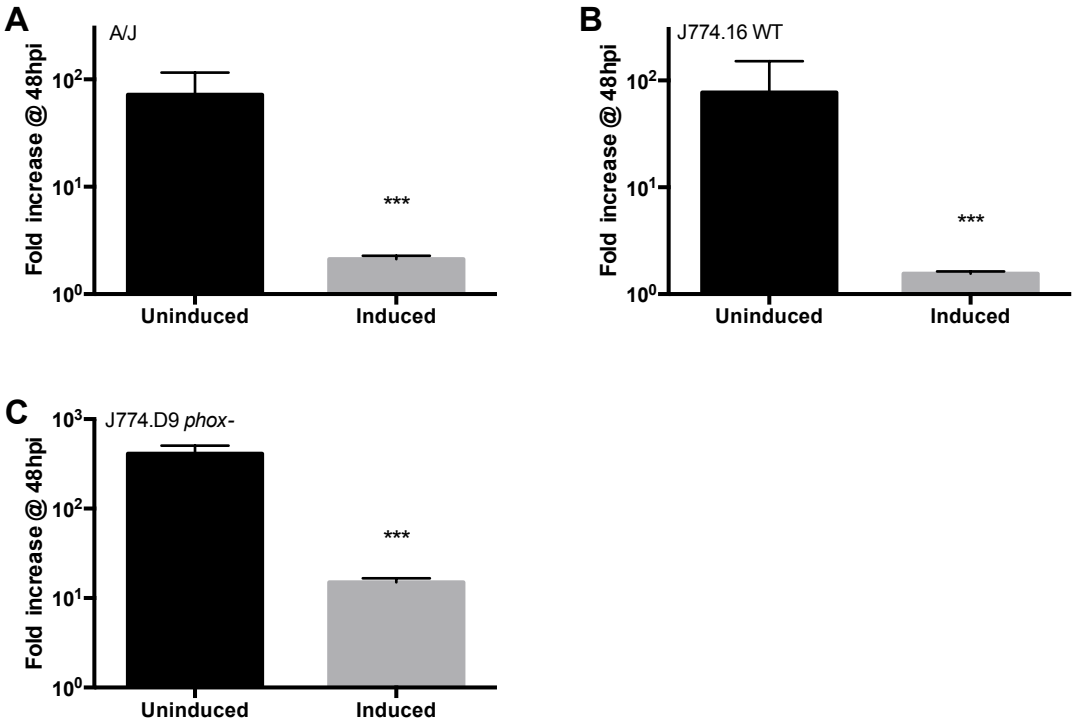

Figure 3.4: CsrT inhibits replication in macrophages independent of NADPH oxidase (A), J774.16 WT macrophage cell line (B) and J774.D9 *phox* mutant macrophage cell line (C). Macrophages were infected with WT Lp02 carrying pcsrT (MB1383) treated +/- IPTG overnight prior to infection at an MOI of 1 (uninduced) or 2 (induced). Results are expressed as fold increase of bacteria at 48 hours post infection (hpi) relative to the 2-hour timepoint (CFU 48hpi / CFU 2hpi) and shown are means +/- SEM from one experiment representative of three others. In all macrophage backgrounds, *t*-tests confirm differences between uninduced and induced strains are significant ($p < 0.005$).

Figure 3.5.

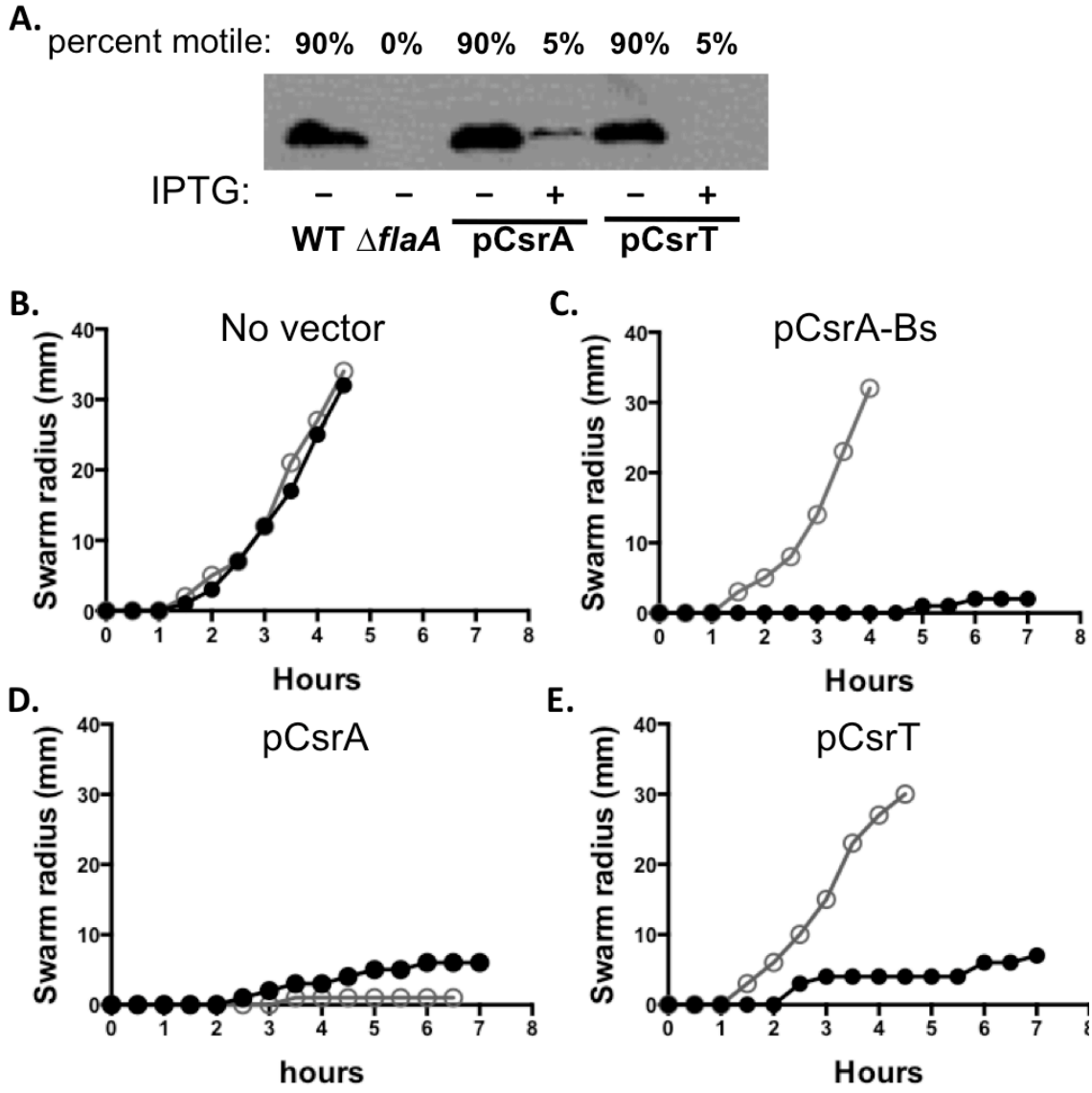

Figure 3.5. CsrT inhibits motility via a conserved mechanism. (A) Western Blot for FlaA protein was performed on PE-phase cultures of Lp02 with pcsrA (MB1389) or pcsrT (MB1383), induced or not as indicated. Lp02 WT with no vector (MB110) was used as a positive control, and a *flaA* mutant (MB1390) as a negative control. Qualitative assessment of the percent motility of the cultures by inverted phase microscopy is indicated above. The blot is representative of at least 3 independent experiments. (B-E) Swarming motility was assessed every 30 minutes on swarm agar plates in *B. subtilis* WT (filled circles) and *fliW sow3* point mutants (open circles), which are insensitive to *B. subtilis* CsrA repression. Motility was assessed in WT (3610) and mutant (DS6530) with (B) no vector, (C) induced expression of *B. subtilis* CsrA (DS4940 and DK1469), (D) induced expression of *L. pneumophila* CsrA

(DK675 and DK1470), and (E) induced expression of CsrT (DK677 and DK1471). Plots represent the means of triplicate plates from one experiment.

CHAPTER IV

CONCLUSION

Introduction

Legionellae can persist in incredibly diverse environments. When these microbes contaminate water that is vaporized, particularly by human machinery in populated areas, they threaten public health. Outbreaks of Legionnaire's disease have been traced back to water sources that vary widely in temperature, stagnation level, and nutrient availability; these include ice machines (Yu et al. 2010), spas (Coetzee et al. 2012), cooling towers (Osawa et al. 2014), puddles (Kanatani et al. 2013), fountains (Haupt et al. 2012), air conditioning vents (Bennett et al. 2014), dental equipment (Chikte et al. 2011), and ponds (Olsen et al. 2010). Legionellae's resilience is also apparent from its extremely broad protist host range (Amaro et al. 2015) and ability to persist for long periods of time within biofilms and as a metabolically repressed viable but nonculturable (VBNC) state (Steinert et al. 1997; Murga et al. 2001; Declerck et al. 2009; Al-Bana et al. 2014).

Coinciding with these diverse habitats, legionellae are notable for having an extremely plastic genome. Only ~70% of genes are conserved between species, a feature thought to contribute to its promiscuity (Miyamoto et al. 2003; Cazalet et al. 2004; Gomez-Valero et al. 2011; O'Connor et al. 2011; Gomez-Valero et al. 2013). While some notable efforts have been undertaken to understand which parts of the

genome are essential for the intracellular life cycle of *L. pneumophila* (Miyamoto et al. 2003; O'Connor et al. 2011; Gomez-Valero et al. 2013), there has been little research on the genetic determinants of legionellae persistence in the environment.

Much of *L. pneumophila* genome plasticity is due to horizontally acquired genetic elements, especially integrative conjugative elements (ICEs). Chapter III demonstrates that every legionellae ICE identified encodes a paralog of *csrA*, the well-characterized essential regulator of the pathogen's intracellular life cycle (Molofsky et al. 2003). In addition, chapter II characterizes a highly conserved *csrA* paralog encoded in the core genome. This thesis proposes and begins to test the hypothesis that the legionellae extensively utilize the regulatory strategy of CsrA by encoding several paralogs in their genomes to orchestrate expression of distinct gene-expression profiles to increase fitness in diverse environments.

The discovery of multiple csrA *paralogs in the* Legionella *pan-genome*

The first dual CsrA system was discovered in *Pseudomonas fluorescens* when I undertook this project, and a second dual system in *Pseudomonas aeruginosa* was published during my thesis work (Reimmann et al. 2005; Marden et al. 2013; Morris et al. 2013). However, while a single *csrA* homolog has been shown or suspected to be in ~75% of all bacterial species (Lucchetti-Miganeh et al. 2008; Papenfort et al. 2010), two paralogs within a single genome had been demonstrated only in the pseudomonads, and more than two paralogs had never been described. Several publications in the *L. pneumophila* literature that surveyed or characterized mobile genetic elements had noted the presence of genes with homology to *csrA* (Segal et al.

1999; Glockner et al. 2008), a pattern that first motivated me to exploit the recent and growing knowledge of the *L. pneumophila* pan-genome for *csrA*-like genes.

I discovered 18 distinct genes with homology to *csrA* in 14 surveyed genomes (Table 2.3). Every genome encodes not only a copy of well-characterized canonical *csrA*, but also an uncharacterized gene, which I named *csrR*. The other 16 *csrA*-like genes were exclusively associated with loci encoding type IV secretion systems (T4SSs). These accessory *csrA* genes are found in anywhere from one to eight of the 14 surveyed genomes. However, it is striking that every genome surveyed has at least four *csrA* paralogs, and some have as many as seven. In Chapter II, I characterized the newly discovered second conserved *csrA* paralog. In Chapter III, I analyzed regulation by one of the T4SS-associated *csrA* paralogs. In this chapter, I will consider both of those stories in the broader context of why the legionellae are seemingly unique in the microbial world by retaining so many *csrA*-like genes.

Csr-on-csr violence is not the complete story of csrR *regulation.*

The paralog *csrR* is highly dissimilar in amino acid sequence from *csrA*. Indeed, in *L. pneumophila* Philadelphia-1, *csrR* is only 28% identical to *csrA*, making it the most divergent of the four *csrA* paralogs in the genome (Fig. 2.1). Yet, despite this divergence, CsrR retains the key residues for RNA binding defined in studies of several other CsrA proteins (Mercante et al. 2006; Schubert et al. 2007; Vakulskas et al. 2015). Moreover, Phyre2 protein-folding software predicts that CsrR retains a structure and binding pockets similar to CsrA (data not shown) (Kelley et al. 2009). Based on likely retention of RNA-binding function and my experiments establishing

that CsrR is not redundant with CsrA, I proposed the following hypothesis: CsrR is an RNA-binding protein whose significant genetic drift from CsrA endowed it with a distinct regulon. If indeed this protein pair orchestrates transitions between incompatible phases of the *L. pneumophila* life cycle, it would logically benefit the bacterium to ensure mutually exclusive expression of CsrA and CsrR.

Interestingly, *L. pneumophila* does not maintain exclusive CsrA and CsrR regulons by controlling transcription. Indeed, *csrR* is highly transcribed in E phase when CsrA is active (Sahr et al. 2012). In E phase, the bacterium relies on CsrA to bind *csrR* mRNA at a GGA motif conserved at the N-terminal end of the coding region (Fig. 2.4). A highly active *csrR* promoter coupled with post-transcriptional repression by CsrA implies that it is advantageous for *L. pneumophila* to maintain an intracellular pool of *csrR* transcripts. However, when post-transcriptional repression was bypassed experimentally by mutating the CsrA binding site on *csrR* mRNA (Fig. 2.6), no phenotypic consequences were detected in laboratory conditions (data not shown). Why then maintain these intracellular pools of *csrR* mRNA? The observation that the *csrR* mRNA encodes a second CsrA binding site that may stabilize this transcript could provide some insight into this paradox (Fig. 2.4 and 2.5). Consider one potential model: abrupt changes that arise during exponential growth, such as premature lysis or exocytosis of the replication vacuole by the host, may require rapid expression of CsrR and repression of CsrA. Such conditions could induce a structural change in the *csrR* mRNA that favors stable binding of CsrA at BS2 of the *csrR* mRNA over the transient interactions of CsrA protein at *csrR* mRNA BS3 that trigger mRNA degradation. If so, the *csrR* mRNA

would mimic the ncRNAs RsmY and RsmZ by binding and competitively inhibiting CsrA; simultaneously, any new *csrR* transcripts would be available for translation. Direct testing of this model first requires identification of culture conditions that relieve CsrA repression of *csrR.* Understanding the functional consequences of the reciprocal expression of CsrA and CsrR is a compelling challenge for future research.

However, CsrA is probably not the only regulator of CsrR. Transcription of *csrR* is repressed in PE phase when CsrA is also repressed, as judged by the decrease in *csrR* promoter activity (Fig. 2.5) and decline in transcript levels (Sahr et al. 2012). Apparently, expression of CsrR is unnecessary or disadvantageous when the bacteria naturally transition into the transmissive phase. This hypothesis is consistent with a life cycle in which lysis from one host cell anticipates the potential encounter and uptake by another host cell, a period when transmissive traits are most beneficial.

Entry into PE phase requires repression of CsrA (Molofsky et al. 2003), and because CsrA represses translation of CsrR (Fig. 2.6), then repression of CsrR in PE phase requires inhibition of transcription, since CsrA is no longer active. To coordinate transcriptional repression of *csrR* with repression of CsrA in PE phase, it would be efficient for *L. pneumophila* to use the same regulatory pathways. For instance, when induced in starvation conditions RelA protein initiates a cascade of regulatory events that together activate the transcription factor LetA, which in turn coordinates CsrA repression (Zusman et al. 2002). Indeed, ectopic expression of RelA is sufficient to significantly reduce *csrR* transcript levels (Dalebroux et al. 2010). A BLAST search for the LetA binding site (AGAAATTTCT) (Sahr et al. 2009)

in the *L. pneumophila* Philadelpia-1 genome reveals at least 20 possible genes directly regulated by LetA (data not shown). Accordingly, transcriptional repressors or activators of *csrR* could be identified by either a targeted search of these candidate genes or an unbiased transposon mutagenesis screen. Understanding the regulatory network that controls CsrR activity promises to expand our knowledge of the *L. pneumophila* life cycle while also providing a model for complex bacterial regulatory circuitries.

Finally, phenotypic analysis of a *csrR* mutant strain revealed that the *csrR* locus contributes to prolonged survival in water (Fig. 2.2). Its mutant phenotype predicts that, in the laboratory conditions applied, at some point *csrR* transcription is activated and protein is translated. To determine the kinetics of CsrR activation, a simple experiment would be to subject a strain carrying the BS2/BS3 translational reporter (Fig. 2.5) to a prolonged incubation in 45˚C water (Fig. 2.2) and to monitor fluorescence. Knowing when CsrR is activated would reveal specific laboratory conditions that recruit the putative CsrR regulon. Transcriptional profiles of cells cultured in activation versus repression conditions could identify both loci affected by CsrR and its specific RNA targets, if indeed CsrR is an RNA binding protein. To identify the kinetics of CsrR activation therefore would be a huge benefit to experimentalists pursuing the biological function of the CsrR protein.

L. pneumophila has two core non-redundant *csrA* genes whose regulatory design is predicted to generate distinct cellular fates. The direct interaction between CsrA and *csrR* is one mechanism of regulation, and it is consistent with the RsmA/RsmF system in *P. aeruginosa* (Marden et al. 2013). However, this dual

regulatory circuitry of *L. pneumophila* is clearly far more complex. My thesis research positions the legionellae field to investigate the mechanisms of environmental persistence by exploiting the tools of genetic regulation. More generally, my graduate work provides the broad CsrA field with a dual *csrA* system to study the mechanisms of fitness advantages of reciprocal regulation.

csrT *represents a new class of* csrA-*paralogs devoted to horizontal gene transmission within legionellae*

In addition to their two core *csrA*-like genes, legionellae can encode from one to several more *csrA* paralogs in horizontally acquired ICEs. Unlike any other bacterial species described thus far, legionellae are remarkable both for encoding so many *csrA* genes and because every ICE within the legionellae encodes a *csrA* paralog.

Why do legionellae ICEs retain *csrA*? For the purpose of this discussion, let us first adopt the perspective that an ICE is a selfish genetic element that ultimately retains only genes beneficial to its maintenance in a host genome and/or its excision, replication, and conjugative spread to new hosts. Accordingly, their ubiquitous presence in these elements indicates that *csrA* genes are beneficial to the ICE. Furthermore, these *csrA* paralogs are genetically linked and apparently co-inherited specifically with the T4SS loci on the ICEs, but not their cargo regions (Fig. 3.1, data not shown). Recall that in the *L. pneumophila* core genome, the canonical *csrA* is a well-characterized pluripotent regulator that, indeed, represses the core Dot/Icm T4SS (Rasis et al. 2009; Nevo et al. 2014). Therefore, a reasonable

hypothesis to account for why legionellae ICEs retain *csrA* paralogs is that this family of regulatory genes regulates appropriate expression of the ICEs' T4SSs.

To test the hypothesis that accessory CsrA proteins regulate their corresponding T4SS, Chapter III examined *csrT* and ICE-βox. The 10-fold reduction in the rate of ICE-βox conjugation (Fig. 3.2) after ectopic expression of *csrT* is consistent with it repressing the ICE-βox T4SS; albeit the inhibition is not absolute. Therefore, ectopic expression of plasmid-born *csrT* in laboratory conditions was not sufficient to establish a functional relationship between CsrT and its associated T4SS. A limitation to our experimental design is that conditions that promote efficient excision or conjugation by ICE-βox T4SS have yet to be identified. We know ICE-βox enhances *L. pneumophila* macrophage infection and hydrogen peroxide resistance (Flynn et al. 2014), and ectopic expression of *csrT* repressed each of these phenotypes (Fig. 3.3 and 3.4). However, whether either phenotype depends on the T4SS apparatus is unknown.

A more direct test of T4SS function in the presence of ectopic *csrT* expression is warranted. For example, it would be straightforward to test the effect of ectopic *csrT* expression on mRNA levels of some of the putative pilus genes encoded on ICE-βox; *lpg2093* and *lpg2087* both have CsrT recognition sequences similar to *hag* in their 5' untranslated regions. Indeed, the fact that *csrT* is encoded directly 5' of the putative pilin gene *lpg2093* (*traA*) and 3' of *lvrB,* which has phyre2 homology to a protein-protein interaction protein suggests analogy to the *B. subtilis* model of flagellar regulation (Fig. 4.1). In *B. subtilis*, CsrA binds the flagellar subunit-encoding *hag* mRNA to repress the flagellum nanomachinery assembly, and FliW binds CsrA

to relieve the repression (Mukherjee et al. 2011; Mukherjee et al. 2013). By analogy, in *L. pneumophila* ICE-βox, CsrT could bind the recognition sequence on the pilin subunit-encoding *traA* mRNA to repress the T4SS pilus nanomachinery, and LvrB could bind CsrT to relieve the repression (Fig. 4.1). This model could explain why there are always *csrA* paralogs immediately 5' of T4SS operons and 3' of an *lvrB* gene in *L. pneumophila* ICEs.

Another informative approach to test if *csrT* affects the function of its T4SS would be to examine the effects of ectopic *csrT* expression on the profile of proteins secreted into the culture supernatant after treatment with hydrogen peroxide. In parallel, one could measure whether culture supernatants collected during *csrT* induction contained less peroxide reducing activity than uninduced control supernatants. Marked differences in the protein profiles and reducing activity not only could link the peroxide resistance conferred by ICE-βox to secreted effectors, but also would support a model in which CsrT inhibits secretion of those effectors.

Although direct binding assays are required to confirm that CsrT protein binds RNA, the functional assays in *B. subtilis* strongly predict that CsrT is an RNA binding protein. First, CsrT retains the key CsrA residues for RNA binding (Mercante et al. 2006; Vakulskas et al. 2015). Second, a point mutation in the *hag* mRNA at the CsrA binding site abrogated motility repression by CsrT (Fig. 3.5). That CsrT's interaction with *hag* was so specific enables bioinformatic predictions about CsrT targets, such as the one proposed above for *traA*. Further, CsrT functional studies in *B. subtilis* validate my use of amino acid sequence, and in particular the

residues that form the RNA-binding pockets, as a bioinformatic probe to identify *csrA* paralogs.

A consideration for future studies

The nuances of local concentration of CsrA protein and RNA targets might be a hurdle to future analysis of ICE-associated *csrA* paralogs. Ectopic expression of several *csrA* genes in their non-native species is sufficient to complement a variety of phenotypes. For instance, *L. pneumophila* CsrA has retained the ability to regulate motility, pigmentation, and cell size of *E. coli* (Fettes et al. 2001). Based on the observation that *L. pneumophila* CsrA can substitute for its counterpart in *E. coli*, we deduce that canonical CsrA of *L. pneumophila* is similar in sequence and structure to the *E. coli* CsrA. When expressed ectopically, the *L. pneumophila* regulator apparently overcomes the finer nuances of concentration-dependent high and low affinity interactions with mRNA targets. CsrA proteins have different affinities for mRNA targets based on sequence of the RNA binding site and side chains of the CsrA protein (Duss et al. 2014). Local concentration of protein and RNA therefore dictate whether lower or higher affinity interactions take place. (Indeed, local concentrations of mRNAs are not constant in the intracellular space of bacteria (Montero Llopis et al. 2010)).

The amount and location of target mRNAs is particularly interesting to consider for CsrT and its associated T4SS. On every *L. pneumophila* ICE, these regulatory and structural genes are adjacent. Because transcription and translation are coupled in bacteria, the concentrations of CsrT protein is predicted to be highest

in the vicinity of the newly synthesized T4SS mRNA. By this logic, when expressed from its native ICE, regulation by CsrT of its T4SS mRNAs may be more efficient, whereas regulation of *L. pneumophila* core genome traits would be less so due to more dilute concentrations of distal mRNAs and CsrT. It is a challenge of any experimental system to mimic native conditions; however, it may be critical to consider chromosomal position during future studies of CsrT regulation of its associated T4SS and cargo loci. If local concentrations of CsrA proteins and their mRNA targets are part of the regulatory mechanism that have evolved to increase fitness of *L. pneumophila* ICEs, ectopic expression may be insufficient to reproduce *bona fide* CsrT interactions with its native targets.

Concluding remarks

Many bacteria use homologs of CsrA as pluripotent regulators of often complex genetic networks to enact widespread changes in protein expression. In my thesis I have established that the legionellae are only the second genus known to encode a second *csrA* locus in their genomes. As the two regulatory genes are not redundant and likely control mutually exclusive phases of the legionellae life cycle, reciprocal expression of CsrA and CsrR is necessary to avoid cellular mayhem. I demonstrate here that CsrA directly binds *csrR* mRNA and represses its translation; however, this direct post-transcriptional regulation is likely just one facet of a more complex regulatory interaction. Furthermore, ICEs unique to the legionellae each encode their own paralog of *csrA*, vastly increasing the number of *csrA* alleles in the legionellae pan-genome. These paralogs in the accessory genome represent an

entirely new class of *csrA*-like genes that appear to be genetically and functionally coupled to their T4SSs. Together, *csrR* and the ICE-associated *csrA* paralogs present a new and rich opportunity to gain greater insight not only to mechanisms of CsrA regulation but also to horizontal transmission of ICEs, the *L. pneumophila* life cycle, and persistence of this opportunistic pathogen in natural and built environments.

Figure 4.1

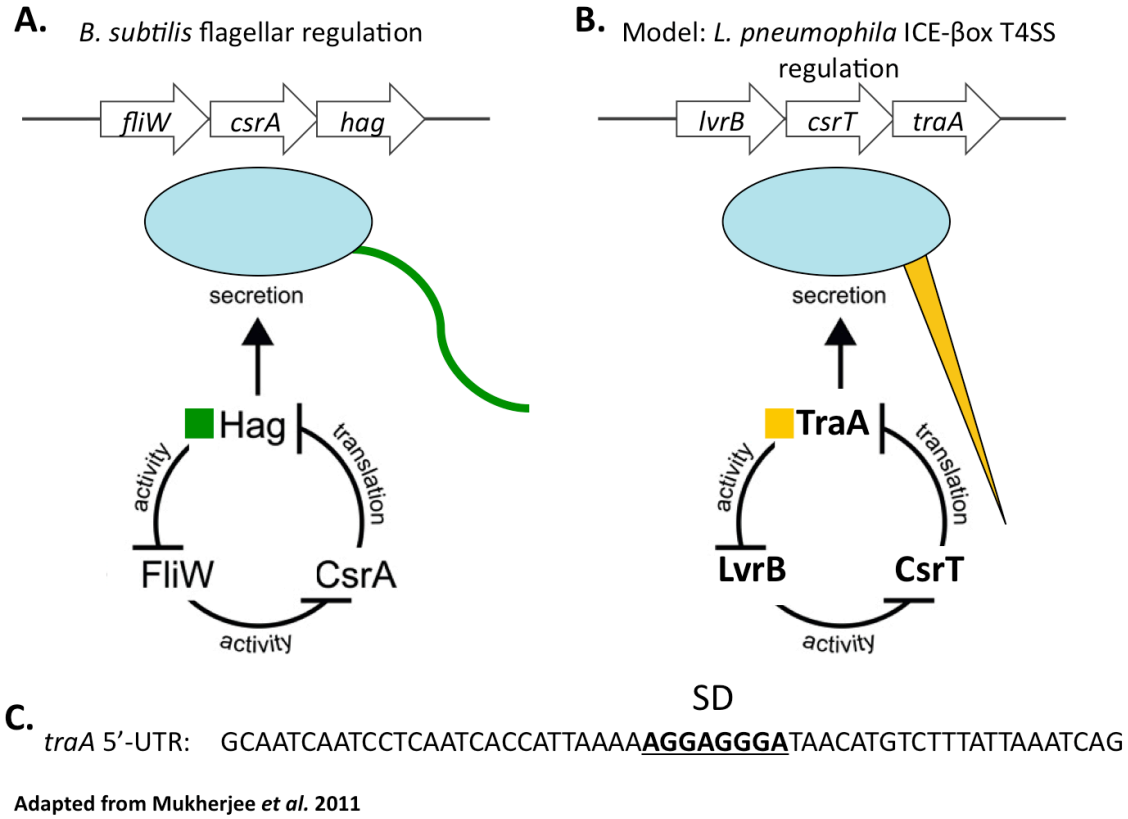

Figure 4.1. Modeling CsrT regulation of ICE-βox T4SS by analogy to *B. subtilis.* (A) In *B. subtilis,* CsrA binds to the 5'-untranslated region (5'-UTR) of the *hag* (flagellin) mRNA and represses translation of flagellin subunits. FliW protein binds to CsrA protein to relieve repression of *hag*, allowing for translation of flagellin and assembly of the flagellum nanomachinery (green structure on blue bacterial cell). Once the flagellum is assembled, excess flagellin protein accumulates in the cytoplasm and binds FliW, thus relieving FliW repression of CsrA. CsrA is then free to bind *hag* mRNA and prevent its translation (Mukherjee et al. 2011; Mukherjee et al. 2013). (B) Analogous to the flagellum system in *B. subtilis*, the *csrA* paralog *csrT* is directly 5' of *traA*, the putative pilin gene of the ICE-βox T4SS (orange structure on blue bacterial cell), which is an evolutionarily related nanomachine to flagellum. In addition, *csrT* is 3' to *lvrB*, a gene predicted to encode a protein with protein-protein interaction function that could enable it to bind to CsrT like FliW to CsrA. (C) Further supporting the model that CsrT represses the ICE-βox T4SS in *L. pneumophila* like CsrA represses flagellum in *B. subtilis* is the putative CsrT binding site overlapping the SD in the *traA* 5'-UTR.

APPENDIX

Constructing Unmarked Gene Deletions in *Legionella pneumophila*

Andrew Bryan, Zachary D. Abbott, and Michele S. Swanson; *methods mol biol* 2013

Volume 954:197-212

Summary

The ability to construct recombinant alleles efficiently in strains of interest, particularly unmarked deletions that reduce the potential for polar effects, is essential to studies of both pathogenesis and basic bacterial physiology. Here we describe a three-phase approach for generating unmarked deletions in *Legionella pneumophila* by constructing a mutant allele in *E. coli* using λ-Red recombination, so-called recombineering; transferring the allele onto the *L. pneumophila* chromosome by natural transformation; and then removing the selectable marker by utilizing the Flp site-specific recombinase. This strategy can decrease the amount of clone screening required while also increasing the percentage of the time the desired allele is obtained on the first attempt. The approach is particularly suited for constructing multiple unmarked deletions in a single strain in fewer steps than traditional methods.

1. Introduction

The three-step method of engineering unmarked deletions in *Legionella pneumophila* involves construction of a recombinant allele, transfer to *L.*

pneumophila, and excision of the resistance cassette (Fig. A.1). Construction of the recombinant allele can be achieved either *in vitro* by standard cloning techniques or *in vivo* by recombineering in *E. coli.* Using primer design and over-lap extension PCR (Ho et al. 1989; Horton et al. 1989; Heckman et al. 2007), alleles can be constructed at the precise nucleotides desired. This technique is also theoretically the fastest method, since recombinant alleles can be constructed by two sequential PCR steps in a single day. A significant drawback, however, is that many experimentalists have difficulty assembling recombinant alleles by PCR reproducibly: Although faster in theory, *in vitro* engineering is often delayed by troubleshooting. Instead, we typically construct recombinant alleles by recombineering in *E. coli*, which utilizes the λ-Red recombinase system to mediate homologous recombination with as few as 36 nucleotides of homology that are specified by primer design (Datsenko et al. 2000; Yu et al. 2000; Court et al. 2002; Thomason et al. 2007; Thomason et al. 2007). A couple of extra days are required first to construct the allele *in vivo* in a recombineering-proficient *E. coli* strain (e.g. DY330), and then to purify the plasmid by transformation of a standard cloning strain (e.g. DH5α). Nevertheless, recombineering has distinct advantages. Not only can the investigator design the allele to the precise nucleotides desired, but recombineering is also extremely efficient. Hundreds or thousands of recombinant clones are often obtained, and, once the method has been established in a laboratory, the desired allele is routinely achieved in the first attempt. Recombineering protocols have been widely published (Thomason et al. 2007;

Thomason et al. 2007; Sharan et al. 2009); Section 3.1 describes one method that works well in our laboratory.

Once a recombinant allele harboring a resistance cassette is constructed, the mutation is transferred to the bacterial chromosome by one of many methods, chiefly natural transformation of *L. pneumophila*. Exploiting natural competence is convenient, but mating or electroporation of recombinant alleles produces the same results. Previous studies of natural transformation describe optimal conditions (Stone et al. 1999; Sexton et al. 2004); Section 3.2 outlines one approach that works particularly well with the hypercompetent strain Lp02. Other laboratory strains, as well as natural isolates, are transformed less efficiently (11; Note 22).

After allele construction, natural transformation, and colony purification, a mutant strain can be used to test hypotheses genetically. However, the residual resistance cassette at the engineered locus can be problematic. The foremost drawback for researchers is the likelihood of polar effects on downstream genes (Bryan et al. 2011). Additionally, in *L. pneumophila* and many other species, few resistance cassettes are available, limiting the number of deletion mutations or plasmids possible for a single strain. Selectable markers flanked by short *FRT* recognition sequences can be excised by the Flp recombinase. This step leaves a small 'scar' sequence from the *FRT* site and priming sequences. A significant advantage of the Flp recombinase is its high activity in a large number of species. When the enzyme is expressed, *FRT*-flanked cassettes are efficiently excised, limiting the number of clones necessary to screen. We recently applied Flp to

generate unmarked mutants in *L. pneumophila* (Bryan et al. 2011); Section 3.3 describes our protocol.

Several alternative approaches to mutant construction are available. The Notes section briefly discusses techniques that utilize over-lap extension PCR, counter-selection (Goodwin et al. 1998; Reyrat et al. 1998; LeBlanc et al. 2006; Bryan et al. 2011), and oligo mutagenesis (Bryan et al. 2011); other options are beyond the scope of this chapter. A traditional approach that has been applied widely and continues to be useful for *L. pneumophila* is allelic exchange using suicide plasmids (Ott 1994), particularly the vector pSR37s (Merriam et al. 1997). Although not the focus of this chapter, allelic exchange continues to be useful. Therefore we provide an additional tool that may be adopted for various uses (see Note 3): pR6K*cat-rdxA-rpsL* encodes a chloramphenicol resistance marker and a double-counter selectable cassette that confers both streptomycin and metroniadazole sensitivity (Bryan et al. 2011). We have constructed another vector that can be used for allelic exchange: pR6K*cat-rdxA-rpsL* encodes a chloramphenicol resistance marker and a double-counter selectable cassette that confers sensitivity to both streptomycin and metroniadazole (Bryan et al. 2011). Since DH5α itself is streptomycin sensitive, and the vector is unstable in the strain, we also generated an *E. coli* strain for amplifying sub-clones and verifying selectable marker phenotypes: HB101 λ*pir* Δ*endA* (strain MB740; see Note 3; 13). However, this cassette is useful only in streptomycin-resistant laboratory strains, such as Lp02, JR32, and AA100; natural isolates of *L. pneumophila* are sensitive to streptomycin.

Although we have primarily used plasmid pR6K*cat-rdxA-rpsL* for other purposes, the vector may be useful for traditional allelic exchange, as the frequency of acquiring spontaneous mutations in both counter-selectable alleles is lower than in any single marker. Allelic exchange using suicide vectors is one useful approach; here we focus on strategies to construct unmarked deletions. By using the Flp site-specific recombinase, one or more resistance cassettes can be removed from the chromosome with high frequency, thus saving time as investigators construct multiple unmarked deletions in a single strain.

2. Materials

2.1 Recombineering components (see Note 1)

1. Template wild-type DNA and primers to amplify 500-1000 bp region surrounding gene to be deleted
2. Thermocycler, reagents, and DNA polymerases for both a high fidelity PCR (e.g. Platinum Taq HiFi, Invitrogen) and economical screening reactions (e.g. Taq DNA polymerase, NEB)
3. Agarose gel electrophoresis equipment and reagents
4. A cloning plasmid that does not replicate in *L. pneumophila* and relevant reagents, or a kit (e.g., pGEM T easy, Promega. Note that although plasmids with colE1 origins are generally unstable in *L. pneumophila,* they do replicate.)
5. Heat- or electro-competent cloning strain (e.g. *E. coli* DH5α)
6. Template plasmid for antibiotic resistance cassette (see Note 3)
    a. pKD3 for chloramphenicol (Datsenko et al. 2000)

b. pKD4 for kanamycin (Datsenko et al. 2000)

c. pR6K*FRT*gent for gentamicin (see Note 3.10; strain MB838, Bryan et al. 2011)

7. Primers to mediate recombineering event (see Note 4)
8. Recombineering strain cultured at 30-32°C (e.g. E. coli DY330; see Note 3.16; Yu et al. 2000)
9. Antibiotics (e.g. ampicillin, chloramphenicol, kanamycin and/or gentamicin)
10. LB broth and LB plates containing relevant antibiotics (ampicillin 100 μg/ml, chloramphenicol 25 μg/ml, kanamycin 50 μg/ml and/or gentamicin 10 μg/ml, streptomycin 1 mg/ml, metronidazole 10 μg/ml)
11. 37°C incubator
12. 30-32°C incubator
13. 30-32°C shaking water bath
14. 42°C shaking water bath
15. Sterile distilled, deionized water (ddH_2O)
16. Electroporator and 1 mm cuvettes
17. Centrifuge and tubes to spin 50 mL at 4600X*g* at 4°C
18. Microcentrifuge and 1.5 mL tubes at 4°C
19. Plasmid preparation reagents (e.g. Qiagen MiniPrep kit)
20. Reagents to purify PCR products (e.g. Qiagen PCR Purification kit and/or Gel Extraction kit)
21. PCR reagents, restriction enzymes, or sequencing reagents to confirm constructs

2.2 Natural transformation components

1. Parent *L. pneumophila* strain (e.g. Lp02)
2. Charcoal-yeast extract plates, supplemented with thymidine if parent strain is an auxotroph (CYE(T))
3. ACES-buffered yeast extract broth, supplemented with thymidine (100 μg/ml) if needed (AYE(T))
4. 37°C incubator and wheel or water bath to incubate test tubes
5. Recombinant antibiotic-resistant allele, circular or linear, generated above (see Note 12)
6. Microcentrifuge and 1.5 mL tubes
7. 28-30°C incubator
8. Sterile PBS
9. Sterile cotton swabs
10. CYE plates supplemented with thymidine (100 μg/ml; CYET) if needed and relevant antibiotic
    (ampicillin 100 μg/ml, chloramphenicol 5 μg/ml, kanamycin 25 μg/ml and/or gentamicin 10 μg/ml, streptomycin 500 μg/ml, metronidazole 10 μg/ml)

2.3 Flp excision components

1. Mutant *L. pneumophila* strain containing *FRT*-flanked cassette
2. pBSFlp (strain MB791, maintain gentamicin selection, Bryan et al. 2011)
3. Sterile ddH_2O
4. Electroporator and 1 mm cuvettes
5. Centrifuge and tubes to spin 50 mL at 4600X*g* at 4°C

6. Microcentrifuge and 1.5 mL tubes at 4°C
7. CYE(T) plates containing 10 μg/mL gentamicin and 200 μM IPTG
8. CYE(T) plates without antibiotics
9. CYE(T) plates containing 5% sucrose (see Note 20)
10. CYE(T) plate containing relevant antibiotics to confirm phenotype
11. PCR and/or sequencing reagents to confirm constructs

3. Methods

3.1 Make recombinant allele by recombineering in *E. coli* (see Note 1)

1. Amplify the gene or region to be deleted and 500-1000 bp of flanking sequences by PCR using high fidelity DNA polymerase and standard techniques.

2. Clone DNA fragment containing gene and flanking DNA into a cloning vector that does not replicate in *L. pneumophila*. We have used the pGEM T easy cloning kit from Promega with ease and success by following the manufacturer's instructions. Prepare plasmid DNA ~3 mL culture, and resuspend plasmid in ddH_2O, not a buffer.

3. For recombineering, design primers to amplify antibiotic resistance cassette from template plasmid. When an *FRT*-flanked cassette is desired, this typically involves using either pKD3 (encoding chloramphenicol resistance), pKD4 (encoding kanamycin resistance), or pR6K*FRT*gent (encoding gentamicin resistance), cassettes that all have the same priming site sequence (see Notes 2-4).

a. P0 oligo

5'-[≥36 nt homology upstream of gene of interest]-

TGTGTAGGCTGGAGCTGCTTC-3'

b. P2 oligo

5'-[≥36 nt homology downstream of gene of interest]—

CATATGAATATCCTCCTTAGTTCC-3'

4. Amplify cassette from template plasmid using high fidelity DNA polymerase. Isolate the DNA product using a PCR purification kit or gel extraction and eluting in ddH2O, not a buffer (see Notes 5, 6). Removing salts is critical for efficient electroporation.

Recommended conditions using Platinum taq HiFi (Invitrogen)

a. reaction

76 µL ddH_2O

4 µL 50mM $MgSO_4$

10uL buffer

2 µL dNTPs (10mM each)

2 µL template plasmid (e.g. pKD3/pKD4/pR6K*FRT*gent)

3 µL of 10 µM working stock of P0 primer (see Note 7)

3 µL of 10 µM working stock of P2 primer (see Note 7)

0.5 µL DNA polymerase

b. incubation cycles

94°C 2 min; (94°C 30 sec; 48°C 30 sec; 68°C 2 min) X 30x; 68°C 7 min

5. Culture *E. coli* DY330 or another recombineering strain overnight at 30-32°C in 2 ml LB broth without antibiotic selection (see Note 8).

6. Transfer 750 µL of overnight DY330 culture into each of two separate 250 mL flasks that contain 50 mL LB (i.e., 1:70 dilutions). Culture at 30-32°C with shaking at 220 r.p.m. until OD_{600} = 0.4-0.5, approximately 3-4 h. The growth phase is critical: Do not exceed OD_{600} = 0.5.

7. When culture density reaches OD_{600} = 0.4-0.5, induce enzyme by transferring one of the flasks to 42°C with shaking for 15 min. Exact temperature and time are critical. Continue to incubate the other flask at 30°C (uninduced cells).

8. After 15 min, immediately place both flasks in an ice water slurry for 5-10 min.

9. Transfer cultures to pre-chilled 4°C 50 mL conical tube and centrifuge for 7 min at 4600X*g* at 4°C.

10. Decant supernatant, resuspend pellets in 35 mL of sterile ice-cold ddH_2O, and then centrifuge for 7 min at 4600X*g* at 4°C.

11. Decant supernatant, resuspend pellet in a total of 1 mL sterile ice-cold ddH_2O, and then transfer to a microcentrifuge tube. Centrifuge at full speed in a 4°C microcentrifuge for 20-60 sec (as short as possible to pellet cells).

12. Wash cells with 1 mL sterile ice-cold ddH_2O. Resuspend cells to a final volume of ~400 µL (see Note 10). If freezing cells, conduct final two washes with sterile ice-cold 10% glycerol.

13. Electroporate 50 µL of cells with DNA mixture using pre-chilled 1 mm cuvettes and settings of 1.8 kV, 200 ohms, and 25 µF. Immediately transfer cells into 1 mL of LB. Do ≥2 electroporations for each desired construct. Volumes of DNA are listed for ease of use, assuming average yield of miniprep plasmid and concentrated PCR product as determined by agarose gel electrophoresis (see Note 5).

    a) Experimental: 50 µL induced cells + 3 µL plasmid miniprep + 7 µL PCR reaction

    b) No enzyme induction control: 50 µL uninduced cells + 3 µL plasmid miniprep + 7 µL PCR reaction

    c) No template DNA control: 50 µL induced cells + 3 µL H20 + 7 µL PCR reaction

    (optional but suggested if high background is observed in uninduced control sample)

    d) No donor DNA control: 50 µL induced cells + 3 µL plasmid + 7 µL

H2O

(optional control, but suggested if high background is observed in uninduced control sample)

14. Incubate cultures for 2-12 h (typically ~3 h) at 30-32°C.

15. Plate cells on selective medium and incubate at 30-32°C for 1-2 days (see Note 11). If more than a few colonies are observed on the uninduced control plate, determine their source by repeating with each of the optional controls.

16. Colony-purify and then screen candidates by PCR using appropriate primers and a small amount of purified colony as a template. Screen ~4-8 clones, as typically >50% of clones have desired allele, but occasionally more screening is required. Note that DY330 transformants often contain a mixed plasmid population. During this preliminary screen, extensive verification is not necessary.

17. Pick 2-3 clones verified to contain desired construct (within a pure or mixed plasmid population). Prepare plasmid DNA, and then transform DH5α or other non-recombineering cloning strain using a dilution of DNA determined empirically to yield <100 colonies per 100 μL of undiluted culture per plate. The goal is to dilute the DNA preparation so that each *E. coli* cell receives only one plasmid molecule. Typically, dilutions of 1:10, 1:100, or 1:1000 are appropriate,

depending on the efficiency of the competent cells and the plasmid concentration.

18. Incubate the transformation mixture overnight at 37°C. Colony purify ~2 DH5α clones from each of the DY330-isolated plasmids electroporated, and then verify the desired construct by PCR, restriction digestion, and/or sequencing. Nearly every clone should contain a homogenous population of the desired plasmid. Next this recombinant allele can be transferred to the *L. pneumophila* chromosome.

3.2 Transfer allele into *L. pneumophila* by natural transformation (see Note 22)

1. Prepare plasmid DNA encoding recombinant allele or use PCR to generate a high concentration of the linear recombinant allele suspended in sterile ddH_2O (see Note 12).

2. Inoculate 5 mL of AYE(T) with parent *L. pneumophila* strain obtained from solid medium and < 1 week old. Culture overnight at 37°C to OD_{600} = 0.7 – 1.2 (see Note 13).

3. Pre-warm CYE(T) plates to 28-30°C.

4. Collect cells from 3 mL of culture by subsequent full speed 1 min centrifugations of two 1.5 mL culture aliquots in one microcentrifuge tube using a

microcentrifuge. Resuspend final cell pellet in 50 µL of AYE(T).

5. Place 25 µL of culture suspension in each of two spots on one pre-warmed CYE(T) plate. Add 25 µL of DNA to one cell sample and 25µL of sterile water to other.

6. Incubate at 30 °C for 2 days, leaving plates upright either until liquid is absorbed or for entire incubation period.

7. Collect each cell patch with a sterile cotton swab and transfer into 1 mL sterile PBS.

8. Plate cell suspension on appropriate selective medium, and incubate 3-5 days at 37°C or longer if mutant grows slowly (see Note 14).

9. Colony purify candidates and confirm chromosomal integration of desired allele by PCR, Southern analysis, and/or sequencing, being careful to rule-out merodiploids if circular DNA was used (see Note 12). If using linear DNA and no colonies were obtained on the No DNA control, nearly every clone will contain the desired allele.

10. You now have a mutant *L. pneumophila* strain. If an unmarked deletion is desired, continue to Section 3.3. If a double mutant is to be constructed, repeat

procedure to introduce second mutation marked by a different resistance cassette (i.e. chloramphenicol and kanamycin from pKD3 and pKD4, respectively) before proceeding. The Flp recombinase (Section 3.3) can efficiently remove multiple cassettes from one strain in a single step.

3.3 Flp-mediated excision of *FRT*-flanked cassettes (see Note 15)

1. Make electro-competent cells of the mutant strain by method similar to that described for *E. coli* (Note 16). Culture mutant *L. pneumophila* strain in 5 mL of AYE(T) overnight. Dilute culture ~ 25- to 100-fold into 35 ml of AYE(T) without antibiotics in a 250 ml flask, and then incubate at 37°C with shaking at 250 r.p.m. Culture until OD_{600} = 1.1 ± 0.3. Chill cultures on ice for ~ 5-10 min, and then centrifuge 30 ml of culture at 4600 X *g* for 7 min at 4°C. Wash cells once with 30 ml ice-cold sterile ddH_2O, and then transfer to a 1.5 ml tube. After centrifugation at full speed for 1 minute at 4°C in a microcentrifuge, wash cells twice with 1 ml of water. Resuspend cells to a final volume of ~ 240 µl. If freezing cells, conduct final two washes with sterile ice-cold 10% glycerol.

2. Electroporate 2 µL of plasmid preparation of pBSFlp with 50 µL of electro-competent cells using 1 mm path length cuvettes, 1.8 kV, 100 ohms, and 25 µF. Transfer immediately to 1 ml of AYE(T) broth.

3. Incubate cells 1-2 h at 37°C, and then plate on CYE(T) medium containing 10 μg/mL gentamicin and 200 μM IPTG.

4. Incubate plates 5-6 days (see Note 17) or longer if slow-growing mutant.

5. Patch fresh single colonies onto plates without any drugs and culture overnight (see Notes 18, 19).

6. Colony purify cells from patches on CYE(T) + 5% sucrose without antibiotics (see Note 20) and culture for 3-5 days or until isolated colonies form (see Note 21).

7. Patch candidates on CYE(T) containing gentamicin and chloramphenicol and/or kanamycin to verify plasmid and marker loss phenotypically. Confirm genotype by PCR and/or sequencing. Typically, 50-90% of clones have lost both the resistance cassette and the plasmid. Flp-mediated excision leaves a ~85 bp scar sequence that typically is non-polar (Datsenko et al. 2000; Bryan et al. 2011).

4. Notes

1. Although compared to traditional methods constructing recombinant alleles by recombineering in *E. coli* can take extra days, for nearly every locus examined we consistently obtain a large number of the desired recombinant clones on the first

attempt. Alternatively, recombinant alleles can be constructed by over-lap extension PCR or restriction enzyme-based cloning using the template plasmids described in this Chapter. Over-lap extension PCR has the potential to be the fastest method of allele construction and is an excellent choice when cloning any gene product that may be toxic to *E. coli.* While we encourage investigators to attempt over-lap extension for cassette construction, in practice some experimentalists find this method to be less reliable.

2. Only use pR6K*FRT*gent if not planning to excise the gentamicin resistance cassette from the chromosome using the pBSFlp vector, since both utilize $gent^R$ as the selectable marker. The pR6K*FRT*gent plasmid is useful when there is no need to construct an unmarked deletion or when using another Flp-encoding plasmid.

3. Recombinant alleles that contain counter-selectable markers but not *FRT*-flanked cassettes can be used to construct clean unmarked deletions (i.e. do not contain a scar sequence) or to insert sequences into the chromosome (i.e. epitope tags) when used as an allelic exchange vector (discussed in Introduction), using oligo mutagenesis (Bryan et al. 2011), or using recombineering in *E. coli* followed by counter-selection in *L. pneumophila*. For the latter, the selectable/double-counter selectable cassette from pR6K*cat-rdxA-rpsL* (strain MB776, maintain chloramphenicol selection) can be amplified using as the forward primer sequence 5′-TGTGACGGAAGATCACTTCG-3′ and the reverse primer sequence 5′-TTAAGCCTTAGGACGCTTCACG-3′ (Bryan et al. 2011). After the mutant allele is constructed by recombineering in DY330 as outlined in the text (or using over-lap extension PCR), use HB101 Δ*endA* (strain MB739)---not DH5α---for amplifying,

maintaining, and screening clones (Bryan et al. 2011). The *cat-rdxA-rpsL* strain is more stable in this HB101 derivative, which also allows phenotypic screening for streptomycin sensitivity (the HB101 parent strain is resistant to streptomycin).

4. Standardizing primer design is important, since it is easy for even the experienced user to mistakenly order the incorrect primers. One strategy is as follows: Use your program of choice to make a sequence file of your target gene region (including the flanking sequences), and then delete the sequence you intend to remove. Open the sequence file of your template plasmid, copy the selectable marker cassette and its corresponding priming sites (i.e. "P0" to "P2"), and then paste this cassette into the file where you have just deleted your gene *in silico*. The newly joined fragments correspond to the recombineering primers. For example, as the forward primer, order the entire forward primer sequence from the template plasmid ("P0") plus at least 36nt from the flanking region, for a total of at least 60 nt (see text and Fig. A.1). Our laboratory orders primers from IDTDNA, which has a size restriction of 60nt for their smallest scale oligonucleotides. Accordingly, it is economical to limit primer length to 60 nt. If you encounter low efficiencies, a modest extension of the flanking region can increase the frequency of the desired recombinant DNA molecule (Yu et al. 2000).

5. The efficiency of the recombineering reaction is highly dependent on DNA concentration. Therefore we recommend using the product from at least a 100 μL PCR reaction of good yield that has been purified using a PCR purification kit and eluted with ddH_2O to a final volume of 30 μL. Because significant amounts of DNA are lost during gel extraction, use gel purification only when your product contains

multiple relatively abundant species, as visualized by standard electrophoresis methods. Ideally, for each electroporation use 100 nmol – 1 μg of PCR product—the more, the better, unless arcing occurs during electroporation.

6. If DNA template for PCR was linear, chromosomal, or encodes an R6Kγ ori (i.e. pKD3, pKD4, pR6K*FRT*gent, pR6K*cat-rdxA-rpsL*), step #4 is sufficient as described. However, if using a plasmid template that can replicate in your recombineering strain to generate false positive colonies that may out-number the desired recombinants, you must remove circular template by either a DpnI digest and/or gel extraction.

7. When primer concentration is > 0.3 μM, an extra species of ~100 bp is generated. Diluting primers can eliminate extra species.

8. In addition to DY330, several other Lambda-Red strains are available from various investigators. If using a plasmid with a rolling circle mechanism of replication, or if plasmid multimers are a problem, use *E. coli* DY331 or another *recA* recombineering strain (Yu et al. 2000).

9. Frozen recombineering-competent (i.e. temperature-induced DY330) cells are stable for several months after preparation, albeit slightly less efficient than fresh cells. Our lab routinely prepares a large batch of cells, aliquots and freezes them, then thaws on ice for use as needed, which greatly reduces the time required to construct a mutant strain. If low efficiencies are a problem, try freshly prepared competent cells.

10. The number of washes necessary can depend on water source, cell handling, and the electroporator. For the highest transformation efficiencies, use the fewest

washes that yield completely de-salted cells. If electroporated cells "arc" or have a low time-constant, try adding an extra wash.

11. Some clones and selectable markers may take two days to form visible colonies (specifically chloramphenicol from pKD3). In contrast, clones that encode the gentamicin-resistance cassette from pR6K*FRT*gent should be colony-purified after 1 day, since background spontaneous mutants appear with longer incubations. If using the *cat-rdxA-rpsL* allele for counter-selection, a moderate growth defect is observed. For strongest counter-selection, pick the smaller, slower-growing colonies—larger colonies may have acquired spontaneous point mutations.

12. Although more efficient for natural transformation of *L. pneumophila,* circular DNA often generates mereodiploid alleles, which can be more frequent than the desired clones and complicate screening. Therefore we typically use linear DNA, which may yield fewer clones, but more frequently yields the desired mutant and reduces the number of colonies needed to screen. Using the flanking primers originally used to clone the gene region, set up at least a 100 µL PCR, purify the product using a column kit, and elute with 30 µL of ddH_2O. The more DNA used, the more desired recombinant colonies obtained: For high efficiency transformation, use no less than 1 µg of DNA, especially if the DNA is linear.

13. Natural transformation requires cells replicating exponentially. Compared to bacteria obtained from a patch on agar (Sexton et al. 2004), broth-grown cells are transformed more efficiently, since an exponential phase culture is more homogeneous than even cells in a young patch. Once the culture begins to transition

into the post-exponential phase, spontaneous mutants are also more frequent (Bryan et al. 2011, unpublished data).

14. If using the *cat-rdxA-rpsL* allele for counter-selection, a moderate growth defect is observed. Pick the smaller, slower growing colonies for strongest counter-selection.

15. Alternative methods to construct clean unmarked deletions and/or insertions that exploit counter-selection and the *cat-rdxA-rpsL* allele are either oligo mutagenesis (Bryan et al. 2011) or natural transformation of a recombinant allele. Theoretically, oligo mutagenesis is the easiest and fastest technique, and we encourage investigators to try it. However, still in its infancy, in its current form the method is somewhat inefficient. Or, the *cat-rdxA-rpsL* cassette inserted at the locus of interest can be replaced with an umarked allele using natural transformation. For this purpose, either by over-lap extension PCR or another method, construct the desired allele and 500-1000 bp of flanking sequence to mediate homologous recombination. Transform the strain carrying your mutation of interest marked by the *cat-rdxA-rpsL* selectable/double-counter-selectable cassette with either the linear fragment or, for greater efficiency, a circular molecule generated by first cloning the fragment into an appropriate vector. Introduce this unmarked allele by natural transformation, and select recombinants on medium containing metroniadazole (10 μg/mL) and streptomycin (500 μg/mL). Make plates fresh and use within one week. Do not leave plates at room temperature overnight prior to use, as counter-selection is less potent, presumably due to drug deterioration. Colony purify the candidates, and confirm allele is as desired.

16. Electro-competent cells can also be prepared from *L. pneumophila* cultured on solid medium. Using cells obtained from either a fresh plate or an exponential phase broth culture, spread the inoculum across an entire plate, being careful not to gouge agar. Incubate 20-36 h, until culture is lightly confluent. Collect cells with a sterile swab, and then suspend in 1 mL of sterile ddH_2O. Wash cells twice with 1 mL of sterile ddH_2O (room temperature is fine). Resuspend cells in 100-200 µL, depending on their density. Electroporate as above. Although significantly faster and less effort, be aware that the plate method yields more spontaneous mutations in various genes (Bryan et al. 2011, unpublished data), likely because they are not a homogenous population of exponential phase cells. Therefore, for most purposes, our laboratory uses the broth method.

17. The pBSFlp plasmid slows the growth of *L. pneumophila* considerably. Compared to the parental strain, strains typically take at least one extra day to yield distinct colonies.

18. Colonies containing pBSFLP and maintained with gentamicin selection do not remain viable for long: Do not wait more than a couple days before transferring candidates to fresh agar.

19. After selecting colonies transformed with pBSFlp by electroporation, we have observed a mixed population of slightly larger colonies among more frequent but smaller colonies that take an extra 1-2 days to appear. Suspicious that larger colonies harbor spontaneous mutants, we have only characterized the smaller colonies. Although we have no evidence that the different colony morphologies

reflect genetic variation, all mutants originally isolated from such small colonies were subsequently verified by molecular and phenotypic tests.

20. Because the pBSFlp plasmid is not stably maintained, counter-selection with sucrose is generally not required, but it can slightly increase the frequency of isolating strains that have lost the plasmid.

21. Strains should now grow at the expected rate, compared to their slow growth when carrying pBSFlp.

22. Transformation on solid medium has worked well for *L. pneumophila* strain Lp02, but this method was not tested extensively with other strains. There is evidence that some strains undergo natural transformation more efficiently in liquid medium, as described by Sexton and Vogel (Sexton et al. 2004). Consistent with Sexton and Vogel, we have obtained low frequencies of transformation with strains Philadeliphia-1 and JR32, even when transforming in liquid medium. For investigators studying strains that do not transform well on solid medium, we outline a protocol for broth cultures here. Using cells obtained from freshly grown agar cultures, inoculate 3-5 mL of AYE(T) and incubate at 37°C to an OD_{600} = 0.7-1.5. Dilute an aliquot of the culture to OD_{600} = 0.05 in 1 mL of AYE(T). Add 1 μg of DNA and incubate at 30°C with shaking for 2-72 h. Two hours is sufficient for transformation of strain Lp02, whereas 24-72 h may be necessary to obtain transformants of more recalcitrant strains. Finally, plate aliquots on solid medium with appropriate selection for transformants, and then culture at 37°C for 2-4 days until colonies appear.

Table A.1. List of bacterial strains and plasmids.

Strain	Genotype or plasmid	Source
E. coli and plasmids		
DH5α (MB120)*	*supE44* Δ*lacU169* (*80 lacZ*Δ*M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Laboratory collection
DY330 (MB737)*	W3110 Δ*lacU169 gal490* λ*cI857* Δ*(cro-bioA)*	(Yu et al. 2000)
DY331 (MB779)*	W3110 Δ*lacU169 srl*::*Tn10* Δ*recA gal490* λ*cI857* Δ(*cro-bioA*)	(Yu et al. 2000)
MB739	HB101 *endA::FRT*	(Bryan et al. 2011)
MB740	HB101λ*pir endA::FRT*	(Bryan et al. 2011)
(MB782)*	pKD3 (*FRT-cat-FRT* allele)	(Datsenko et al. 2000)
(MB783)*	pKD4 (*FRT-kan-FRT* allele)	(Datsenko et al. 2000)
MB838	DH5α λ*pir endA::FRT* pR6Kγ*FRTgent*	(Bryan et al. 2011)
MB776	HB101λ*pir endA::FRT* pR6K*cat-rdxA-rpsL*	(Bryan et al. 2011)
MB791	DH5α pBSFlp	(Bryan et al. 2011)
L. pneumophila		
(MB110)*	Lp02 wild type, *thyA, hsdR, rpsL* (Str^R)	(Berger et al. 1993)

*Strain number does not correspond to original source; rather it is the strain available from an alternative source, the M.S. Swanson strain collection (MB collection).

Figure A.1.

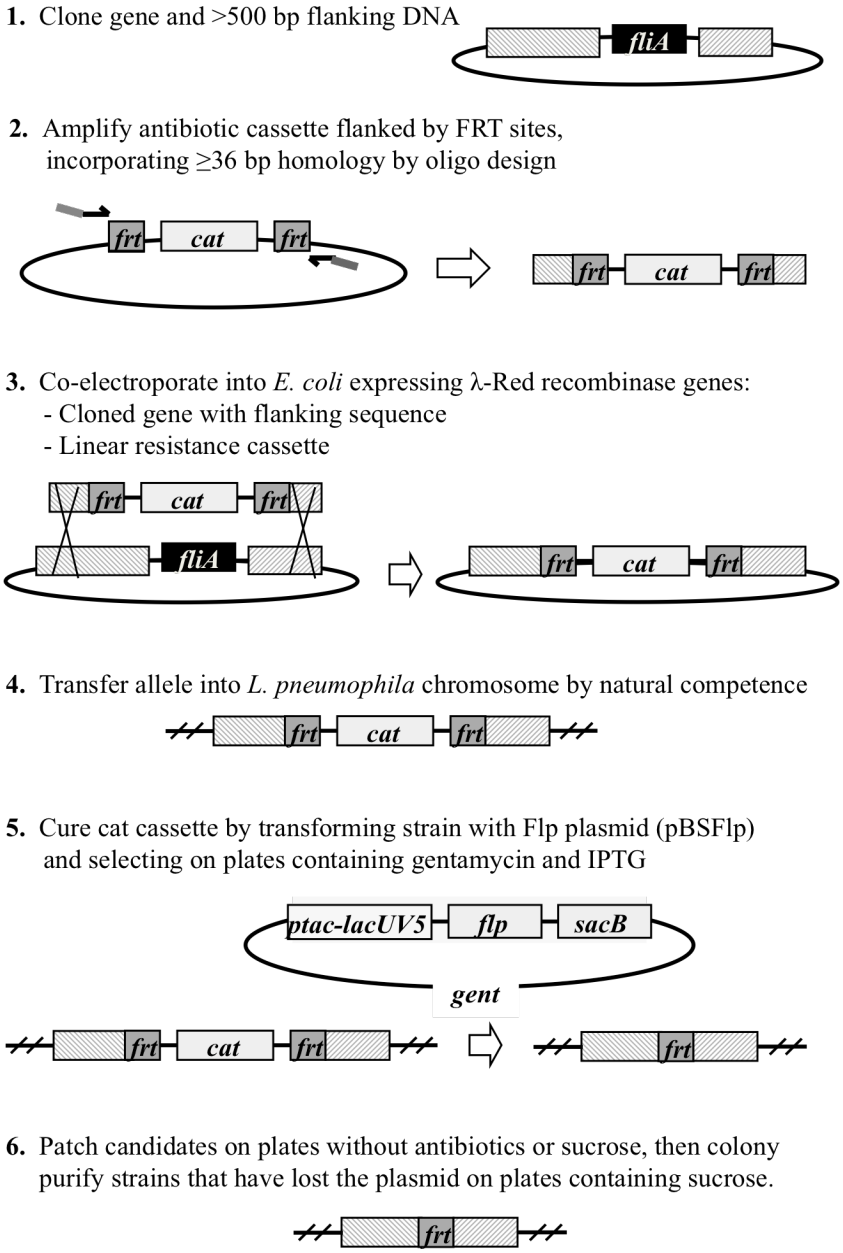

Figure A.1. Outline of strategy for unmarked gene deletions. Outline of strategy with an example of construction of an unmarked deletion in *fliA* using the *FRT*-flanked chloramphenicol resistance cassette from pKD3 followed by excision of the cassette promoted by pBSFlp**.**

BIBLIOGRAPHY

Al-Bana, B. H., M. T. Haddad, et al. (2014). "Stationary phase and mature infectious forms of *Legionella pneumophila* produce distinct viable but non-culturable cells." Environ Microbiol **16**(2): 382-395.

Amaro, F., W. Wang, et al. (2015). "Diverse protist grazers select for virulence-related traits in *Legionella*." ISME J.

Andrade, M. O., C. S. Farah, et al. (2014). "The post-transcriptional regulator *rsmA/csrA* activates T3SS by stabilizing the 5' UTR of *hrpG*, the master regulator of *hrp/hrc* genes, in *Xanthomonas*." PLoS Pathog **10**(2): e1003945.

Baker, C. S., I. Morozov, et al. (2002). "CsrA regulates glycogen biosynthesis by preventing translation of *glgC* in *Escherichia coli.*" Mol Microbiol **44**(6): 1599-1610.

Bandyopadhyay, P., S. Liu, et al. (2007). "Environmental mimics and the Lvh type IVA secretion system contribute to virulence-related phenotypes of *Legionella pneumophila*." Infect Immun **75**(2): 723-735.

Beaber, J. W., B. Hochhut, et al. (2004). "SOS response promotes horizontal dissemination of antibiotic resistance genes." Nature **427**(6969): 72-74.

Bennett, E., M. Ashton, et al. (2014). "Barrow-in-Furness: a large community legionellosis outbreak in the UK." Epidemiol Infect **142**(8): 1763-1777.

Berger, K. H. and R. R. Isberg (1993). "Two distinct defects in intracellular growth complemented by a single genetic locus in *Legionella pneumophila*." Mol Microbiol **7**(1): 7-19.

Berger, K. H., J. J. Merriam, et al. (1994). "Altered intracellular targeting properties associated with mutations in the *Legionella pneumophila* dotA gene." Mol Microbiol **14**(4): 809-822.

Bevilacqua, J. M. and P. C. Bevilacqua (1998). "Thermodynamic analysis of an RNA combinatorial library contained in a short hairpin." Biochemistry **37**(45): 15877-15884.

Bhatt, S., A. N. Edwards, et al. (2009). "The RNA binding protein CsrA is a pleiotropic regulator of the locus of enterocyte effacement pathogenicity island of enteropathogenic *Escherichia coli*." Infect Immun **77**(9): 3552-3568.

Bordi, C., M. C. Lamy, et al. (2010). "Regulatory RNAs and the HptB/RetS signalling pathways fine-tune *Pseudomonas aeruginosa* pathogenesis." Mol Microbiol **76**(6): 1427-1443.

Bose, B., J. M. Auchtung, et al. (2008). "A conserved anti-repressor controls horizontal gene transfer by proteolysis." Mol Microbiol **70**(3): 570-582.

Bose, B. and A. D. Grossman (2011). "Regulation of horizontal gene transfer in *Bacillus subtilis* by activation of a conserved site-specific protease." J Bacteriol **193**(1): 22-29.

Brencic, A. and S. Lory (2009). "Determination of the regulon and identification of novel mRNA targets of *Pseudomonas aeruginosa* RsmA." Mol Microbiol **72**(3): 612-632.
Bryan, A., K. Harada, et al. (2011). "Efficient generation of unmarked deletions in *Legionella pneumophila*." Appl Environ Microbiol **77**(7): 2545-2548.
Bryan, A. and M. S. Swanson (2011). "Oligonucleotides stimulate genomic alterations of *Legionella pneumophila*." Mol Microbiol **80**(1): 231-247.
Byrne, B. and M. S. Swanson (1998). "Expression of *Legionella pneumophila* virulence traits in response to growth conditions." Infect Immun **66**(7): 3029-3034.
Cazalet, C., S. Jarraud, et al. (2008). "Multigenome analysis identifies a worldwide distributed epidemic *Legionella pneumophila* clone that emerged within a highly diverse species." Genome Res **18**(3): 431-441.
Cazalet, C., C. Rusniok, 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.
Chien, M., I. Morozova, et al. (2004). "The genomic sequence of the accidental pathogen *Legionella pneumophila*." Science **305**(5692): 1966-1968.
Chikte, U. M., O. Khondowe, et al. (2011). "A case study of a dental receptionist diagnosed with Legionnaires' disease." SADJ **66**(6): 284-287.
Clewell, D. B. (2011). "Tales of conjugation and sex pheromones: A plasmid and enterococcal odyssey." Mob Genet Elements **1**(1): 38-54.
Coetzee, N., H. Duggal, et al. (2012). "An outbreak of Legionnaires' disease associated with a display spa pool in retail premises, Stoke-on-Trent, United Kingdom, July 2012." Euro Surveill **17**(37).
Cohn, P. D., J. A. Gleason, et al. (2014). "Community outbreak of legionellosis and an environmental investigation into a community water system." Epidemiol Infect: 1-10.
Court, D. L., J. A. Sawitzke, et al. (2002). "Genetic engineering using homologous recombination." Annu Rev Genet **36**: 361-388.
Dalebroux, Z. D., R. L. Edwards, et al. (2009). "SpoT governs *Legionella pneumophila* differentiation in host macrophages." Mol Microbiol **71**(3): 640-658.
Dalebroux, Z. D., S. L. Svensson, et al. (2010). "ppGpp conjures bacterial virulence." Microbiol Mol Biol Rev **74**(2): 171-199.
Dalebroux, Z. D., B. F. Yagi, et al. (2010). "Distinct roles of ppGpp and DksA in *Legionella pneumophila* differentiation." Mol Microbiol **76**(1): 200-219.
Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products." Proc Natl Acad Sci U S A **97**(12): 6640-6645.
Declerck, P., J. Behets, et al. (2009). "Replication of *Legionella pneumophila* in biofilms of water distribution pipes." Microbiol Res **164**(6): 593-603.
Doleans-Jordheim, A., M. Akermi, et al. (2006). "Growth-phase-dependent mobility of the *lvh*-encoding region in *Legionella pneumophila* strain Paris." Microbiology **152**(Pt 12): 3561-3568.
Dubey, A. K., C. S. Baker, et al. (2005). "RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction." RNA **11**(10): 1579-1587.

Dubey, A. K., C. S. Baker, et al. (2003). "CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript." J Bacteriol **185**(15): 4450-4460.
Duss, O., E. Michel, et al. (2014). "Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition." Nucleic Acids Res **42**(8): 5332-5346.
Edwards, A. N., L. M. Patterson-Fortin, et al. (2011). "Circuitry linking the Csr and stringent response global regulatory systems." Mol Microbiol **80**(6): 1561-1580.
Edwards, R. L., Z. D. Dalebroux, et al. (2009). "*Legionella pneumophila* couples fatty acid flux to microbial differentiation and virulence." Mol Microbiol **71**(5): 1190-1204.
Edwards, R. L., M. Jules, et al. (2010). "The *Legionella pneumophila* LetA/LetS two-component system exhibits rheostat-like behavior." Infect Immun **78**(6): 2571-2583.
Faucher, S. P., C. A. Mueller, et al. (2011). "*Legionella Pneumophila* Transcriptome during Intracellular Multiplication in Human Macrophages." Front Microbiol **2**: 60.
Feeley, J. C., R. J. Gibson, et al. (1979). "Charcoal-yeast extract agar: primary isolation medium for *Legionella pneumophila*." J Clin Microbiol **10**(4): 437-441.
Fettes, P. S., V. Forsbach-Birk, et al. (2001). "Overexpresssion of a *Legionella pneumophila* homologue of the *E. coli* regulator *csrA* affects cell size, flagellation, and pigmentation." Int J Med Microbiol **291**(5): 353-360.
Fields, B. S. (1996). "The molecular ecology of legionellae." Trends Microbiol **4**(7): 286-290.
Flynn, K. J. and M. S. Swanson (2014). "Integrative conjugative element ICE-betaox confers oxidative stress resistance to *Legionella pneumophila in vitro* and in macrophages." MBio **5**(3): e01091-01014.
Forsbach-Birk, V., T. McNealy, et al. (2004). "Reduced expression of the global regulator protein CsrA in *Legionella pneumophila* affects virulence-associated regulators and growth in *Acanthamoeba castellanii*." Int J Med Microbiol **294**(1): 15-25.
Frost, L. S. and G. Koraimann (2010). "Regulation of bacterial conjugation: balancing opportunity with adversity." Future Microbiol **5**(7): 1057-1071.
Garduno, R. A., E. Garduno, et al. (2002). "Intracellular growth of *Legionella pneumophila* gives rise to a differentiated form dissimilar to stationary-phase forms." Infect Immun **70**(11): 6273-6283.
Giao, M. S., S. A. Wilks, et al. (2009). "Validation of SYTO 9/propidium iodide uptake for rapid detection of viable but noncultivable *Legionella pneumophila*." Microb Ecol **58**(1): 56-62.
Gibson, D. G., L. Young, et al. (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases." Nat Methods **6**(5): 343-345.
Glockner, G., C. Albert-Weissenberger, et al. (2008). "Identification and characterization of a new conjugation/type IVA secretion system (*trb/tra*) of *Legionella pneumophila* Corby localized on two mobile genomic islands." Int J Med Microbiol **298**(5-6): 411-428.

Goldberg, M., L. S. Belkowski, et al. (1990). "Regulation of macrophage function by interferon-gamma. Somatic cell genetic approaches in murine macrophage cell lines to mechanisms of growth inhibition, the oxidative burst, and expression of the chronic granulomatous disease gene." J Clin Invest **85**(2): 563-569.

Gomez-Valero, L. and C. Buchrieser (2013). "Genome dynamics in *Legionella*: the basis of versatility and adaptation to intracellular replication." Cold Spring Harb Perspect Med **3**(6).

Gomez-Valero, L., C. Rusniok, et al. (2011). "Extensive recombination events and horizontal gene transfer shaped the *Legionella pneumophila* genomes." BMC Genomics **12**: 536.

Gomez-Valero, L., C. Rusniok, et al. (2014). "Comparative analyses of *Legionella* species identifies genetic features of strains causing Legionnaires' disease." Genome Biol **15**(11): 505.

Goodwin, A., D. Kersulyte, et al. (1998). "Metronidazole resistance in *Helicobacter pylori* is due to null mutations in a gene (*rdxA*) that encodes an oxygen-insensitive NADPH nitroreductase." Mol Microbiol **28**(2): 383-393.

Guindon, S., J. F. Dufayard, et al. (2010). "New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0." Syst Biol **59**(3): 307-321.

Gutierrez, P., Y. Li, et al. (2005). "Solution structure of the carbon storage regulator protein CsrA from *Escherichia coli.*" J Bacteriol **187**(10): 3496-3501.

Hammer, B. K. and M. S. Swanson (1999). "Co-ordination of *Legionella pneumophila* virulence with entry into stationary phase by ppGpp." Mol Microbiol **33**(4): 721-731.

Hammer, B. K., E. S. Tateda, et al. (2002). "A two-component regulator induces the transmission phenotype of stationary-phase *Legionella pneumophila*." Mol Microbiol **44**(1): 107-118.

Haudecoeur, E. and D. Faure (2010). "A fine control of quorum-sensing communication in *Agrobacterium tumefaciens*." Commun Integr Biol **3**(2): 84-88.

Haupt, T. E., R. T. Heffernan, et al. (2012). "An outbreak of Legionnaires disease associated with a decorative water wall fountain in a hospital." Infect Control Hosp Epidemiol **33**(2): 185-191.

Heckman, K. L. and L. R. Pease (2007). "Gene splicing and mutagenesis by PCR-driven overlap extension." Nat Protoc **2**(4): 924-932.

Heeb, S., C. Blumer, et al. (2002). "Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0." J Bacteriol **184**(4): 1046-1056.

Ho, S. N., H. D. Hunt, et al. (1989). "Site-directed mutagenesis by overlap extension using the polymerase chain reaction." Gene **77**(1): 51-59.

Horton, R. M., H. D. Hunt, et al. (1989). "Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension." Gene **77**(1): 61-68.

Horwitz, M. A. and S. C. Silverstein (1980). "Legionnaires' disease bacterium (*Legionella pneumophila*) multiples intracellularly in human monocytes." J Clin Invest **66**(3): 441-450.

Ilangovan, A., S. Connery, et al. (2015). "Structural biology of the Gram-negative bacterial conjugation systems." Trends Microbiol.

Jonas, K., A. N. Edwards, et al. (2010). "Complex regulatory network encompassing the Csr, c-di-GMP and motility systems of *Salmonella* Typhimurium." Environ Microbiol **12**(2): 524-540.

Juhas, M., D. W. Crook, et al. (2008). "Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence." Cell Microbiol **10**(12): 2377-2386.

Kanatani, J., J. Isobe, et al. (2013). "Close genetic relationship between *Legionella pneumophila* serogroup 1 isolates from sputum specimens and puddles on roads, as determined by sequence-based typing." Appl Environ Microbiol **79**(13): 3959-3966.

Kao, C. Y., B. S. Sheu, et al. (2014). "CsrA regulates *Helicobacter pylori* J99 motility and adhesion by controlling flagella formation." Helicobacter **19**(6): 443-454.

Kelley, L. A. and M. J. Sternberg (2009). "Protein structure prediction on the Web: a case study using the Phyre server." Nat Protoc **4**(3): 363-371.

Kim, C., S. Jeon, et al. (2015). "Isolation of *Legionella pneumophila* from cooling towers, public baths, hospitals, and fountains in Seoul, Korea, from 2010 to 2012." J Environ Health **77**(6): 58-62.

Klebanoff, S. J. (2005). "Myeloperoxidase: friend and foe." J Leukoc Biol **77**(5): 598-625.

Koraimann, G. and M. A. Wagner (2014). "Social behavior and decision making in bacterial conjugation." Front Cell Infect Microbiol **4**: 54.

Lapouge, K., R. Perozzo, et al. (2013). "RNA pentaloop structures as effective targets of regulators belonging to the RsmA/CsrA protein family." RNA Biol **10**(6): 1031-1041.

Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and Clustal X version 2.0." Bioinformatics **23**(21): 2947-2948.

Lautner, M., E. Schunder, et al. (2013). "Regulation, integrase-dependent excision, and horizontal transfer of genomic islands in *Legionella pneumophila*." J Bacteriol **195**(7): 1583-1597.

Lawhon, S. D., J. G. Frye, et al. (2003). "Global regulation by CsrA in *Salmonella* Typhimurium." Mol Microbiol **48**(6): 1633-1645.

LeBlanc, J. J., R. J. Davidson, et al. (2006). "Compensatory functions of two alkyl hydroperoxide reductases in the oxidative defense system of *Legionella pneumophila*." J Bacteriol **188**(17): 6235-6244.

Letunic, I. and P. Bork (2011). "Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy." Nucleic Acids Res **39**(Web Server issue): W475-478.

Liaw, S. J., H. C. Lai, et al. (2003). "Role of RsmA in the regulation of swarming motility and virulence factor expression in *Proteus mirabilis*." J Med Microbiol **52**(Pt 1): 19-28.

Liu, M. Y., H. Yang, et al. (1995). "The product of the pleiotropic *Escherichia coli* gene *csrA* modulates glycogen biosynthesis via effects on mRNA stability." J Bacteriol **177**(10): 2663-2672.

Liu, Y., P. Gao, et al. (2008). "An *in vivo* gene deletion system for determining temporal requirement of bacterial virulence factors." Proc Natl Acad Sci U S A **105**(27): 9385-9390.
Lucchetti-Miganeh, C., E. Burrowes, et al. (2008). "The post-transcriptional regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages of infection in animal hosts." Microbiology **154**(Pt 1): 16-29.
Marden, J. N., M. R. Diaz, et al. (2013). "An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in *Pseudomonas aeruginosa*." Proc Natl Acad Sci U S A **110**(37): 15055-15060.
Martinez, E., B. Bartolome, et al. (1988). "pACYC184-derived cloning vectors containing the multiple cloning site and *lacZ* alpha reporter gene of pUC8/9 and pUC18/19 plasmids." Gene **68**(1): 159-162.
Martinez, L. C., H. Yakhnin, et al. (2011). "Integration of a complex regulatory cascade involving the SirA/BarA and Csr global regulatory systems that controls expression of the *Salmonella* SPI-1 and SPI-2 virulence regulons through HilD." Mol Microbiol **80**(6): 1637-1656.
Mercante, J., A. N. Edwards, et al. (2009). "Molecular geometry of CsrA (RsmA) binding to RNA and its implications for regulated expression." J Mol Biol **392**(2): 511-528.
Mercante, J., K. Suzuki, et al. (2006). "Comprehensive alanine-scanning mutagenesis of *Escherichia coli* CsrA defines two subdomains of critical functional importance." J Biol Chem **281**(42): 31832-31842.
Merriam, J. J., R. Mathur, et al. (1997). "Analysis of the *Legionella pneumophila* fliI gene: intracellular growth of a defined mutant defective for flagellum biosynthesis." Infect Immun **65**(6): 2497-2501.
Miyamoto, H., S. Yoshida, et al. (2003). "Virulence conversion of *Legionella pneumophila* by conjugal transfer of chromosomal DNA." J Bacteriol **185**(22): 6712-6718.
Molofsky, A. B., L. M. Shetron-Rama, et al. (2005). "Components of the *Legionella pneumophila* flagellar regulon contribute to multiple virulence traits, including lysosome avoidance and macrophage death." Infect Immun **73**(9): 5720-5734.
Molofsky, A. B. and M. S. Swanson (2003). "*Legionella pneumophila* CsrA is a pivotal repressor of transmission traits and activator of replication." Mol Microbiol **50**(2): 445-461.
Montero Llopis, P., A. F. Jackson, et al. (2010). "Spatial organization of the flow of genetic information in bacteria." Nature **466**(7302): 77-81.
Morales, V. M., A. Backman, et al. (1991). "A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants." Gene **97**(1): 39-47.
Morris, E. R., G. Hall, et al. (2013). "Structural rearrangement in an RsmA/CsrA ortholog of *Pseudomonas aeruginosa* creates a dimeric RNA-binding protein, RsmN." Structure **21**(9): 1659-1671.
Mukherjee, A., Y. Cui, et al. (1996). "Global regulation in *Erwinia* species by *Erwinia carotovora rsmA,* a homologue of *Escherichia coli csrA:* repression of

secondary metabolites, pathogenicity and hypersensitive reaction." Microbiology **142 (Pt 2)**: 427-434.
Mukherjee, S., P. Babitzke, et al. (2013). "FliW and FliS function independently to control cytoplasmic flagellin levels in *Bacillus subtilis*." J Bacteriol **195**(2): 297-306.
Mukherjee, S., A. C. Bree, et al. (2015). "Adaptor-mediated Lon proteolysis restricts *Bacillus subtilis* hyperflagellation." Proc Natl Acad Sci U S A **112**(1): 250-255.
Mukherjee, S., H. Yakhnin, et al. (2011). "CsrA-FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in *Bacillus subtilis*." Mol Microbiol **82**(2): 447-461.
Murga, R., T. S. Forster, et al. (2001). "Role of biofilms in the survival of *Legionella pneumophila* in a model potable-water system." Microbiology **147**(Pt 11): 3121-3126.
Nevo, O., T. Zusman, et al. (2014). "Identification of *Legionella pneumophila* effectors regulated by the LetAS-RsmYZ-CsrA regulatory cascade, many of which modulate vesicular trafficking." J Bacteriol **196**(3): 681-692.
O'Connor, T. J., Y. Adepoju, et al. (2011). "Minimization of the *Legionella pneumophila* genome reveals chromosomal regions involved in host range expansion." Proc Natl Acad Sci U S A **108**(36): 14733-14740.
Ohno, A., N. Kato, et al. (2008). "Temperature-dependent parasitic relationship between *Legionella pneumophila* and a free-living amoeba (*Acanthamoeba castellanii*)." Appl Environ Microbiol **74**(14): 4585-4588.
Olsen, J. S., T. Aarskaug, et al. (2010). "Alternative routes for dissemination of *Legionella pneumophila* causing three outbreaks in Norway." Environ Sci Technol **44**(22): 8712-8717.
Osawa, K., K. Shigemura, et al. (2014). "A case of nosocomial *Legionella* pneumonia associated with a contaminated hospital cooling tower." J Infect Chemother **20**(1): 68-70.
Ott, M. (1994). "Genetic approaches to study *Legionella pneumophila* pathogenicity." FEMS Microbiol Rev **14**(2): 161-176.
Pannuri, A., H. Yakhnin, et al. (2012). "Translational repression of NhaR, a novel pathway for multi-tier regulation of biofilm circuitry by CsrA." J Bacteriol **194**(1): 79-89.
Papenfort, K. and J. Vogel (2010). "Regulatory RNA in bacterial pathogens." Cell Host Microbe **8**(1): 116-127.
Patterson-Fortin, L. M., C. A. Vakulskas, et al. (2013). "Dual posttranscriptional regulation via a cofactor-responsive mRNA leader." J Mol Biol **425**(19): 3662-3677.
Pessi, G., F. Williams, et al. (2001). "The global posttranscriptional regulator RsmA modulates production of virulence determinants and N-acylhomoserine lactones in *Pseudomonas aeruginosa*." J Bacteriol **183**(22): 6676-6683.
Qin, T., H. Zhou, et al. (2014). "Distribution of sequence-based types of *Legionella pneumophila* serogroup 1 strains isolated from cooling towers, hot springs, and potable water systems in China." Appl Environ Microbiol **80**(7): 2150-2157.

Qiu, J. and Z. Q. Luo (2013). "Effector translocation by the *Legionella* Dot/Icm type IV secretion system." Curr Top Microbiol Immunol **376**: 103-115.
Rao, C., H. Benhabib, et al. (2013). "Phylogenetic reconstruction of the *Legionella pneumophila* Philadelphia-1 laboratory strains through comparative genomics." PLoS One **8**(5): e64129.
Rasis, M. and G. Segal (2009). "The LetA-RsmYZ-CsrA regulatory cascade, together with RpoS and PmrA, post-transcriptionally regulates stationary phase activation of *Legionella pneumophila* Icm/Dot effectors." Mol Microbiol **72**(4): 995-1010.
Reimmann, C., C. Valverde, et al. (2005). "Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0." J Bacteriol **187**(1): 276-285.
Reyrat, J. M., V. Pelicic, et al. (1998). "Counterselectable markers: untapped tools for bacterial genetics and pathogenesis." Infect Immun **66**(9): 4011-4017.
Rodgers, F. G. (1983). "The role of structure and invasiveness on the pathogenicity of *Legionella*." Zentralbl Bakteriol Mikrobiol Hyg A **255**(1): 138-144.
Rodriguez-Blanco, A., M. L. Lemos, et al. (2012). "Integrating conjugative elements as vectors of antibiotic, mercury, and quaternary ammonium compound resistance in marine aquaculture environments." Antimicrob Agents Chemother **56**(5): 2619-2626.
Rowbotham, T. J. (1980). "Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae." J Clin Pathol **33**(12): 1179-1183.
Roy, C. R. (2002). "The Dot/Icm transporter of *Legionella pneumophila:* a bacterial conductor of vesicle trafficking that orchestrates the establishment of a replicative organelle in eukaryotic hosts." Int J Med Microbiol **291**(6-7): 463-467.
Sahr, T., H. Bruggemann, et al. (2009). "Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*." Mol Microbiol **72**(3): 741-762.
Sahr, T., C. Rusniok, 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.
Schubert, M., K. Lapouge, et al. (2007). "Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA." Nat Struct Mol Biol **14**(9): 807-813.
Segal, G., J. J. Russo, et al. (1999). "Relationships between a new type IV secretion system and the *icm/dot* virulence system of *Legionella pneumophila.*" Mol Microbiol **34**(4): 799-809.
Segal, G. and H. A. Shuman (1998). "Intracellular multiplication and human macrophage killing by *Legionella pneumophila* are inhibited by conjugal components of IncQ plasmid RSF1010." Mol Microbiol **30**(1): 197-208.
Sentchilo, V., R. Ravatn, et al. (2003). "Unusual integrase gene expression on the *clc* genomic island in *Pseudomonas* sp. strain B13." J Bacteriol **185**(15): 4530-4538.

Sexton, J. A. and J. P. Vogel (2004). "Regulation of hypercompetence in *Legionella pneumophila*." J Bacteriol **186**(12): 3814-3825.
Sharan, S. K., L. C. Thomason, et al. (2009). "Recombineering: a homologous recombination-based method of genetic engineering." Nat Protoc **4**(2): 206-223.
Steinert, M., L. Emody, et al. (1997). "Resuscitation of viable but nonculturable *Legionella pneumophila* Philadelphia JR32 by *Acanthamoeba castellanii*." Appl Environ Microbiol **63**(5): 2047-2053.
Stone, B. J. and Y. Abu Kwaik (1998). "Expression of multiple pili by *Legionella pneumophila*: identification and characterization of a type IV pilin gene and its role in adherence to mammalian and protozoan cells." Infect Immun **66**(4): 1768-1775.
Stone, B. J. and Y. A. Kwaik (1999). "Natural competence for DNA transformation by *Legionella pneumophila* and its association with expression of type IV pili." J Bacteriol **181**(5): 1395-1402.
Sullivan, M. J., N. K. Petty, et al. (2011). "Easyfig: a genome comparison visualizer." Bioinformatics **27**(7): 1009-1010.
Suzuki, K., X. Wang, et al. (2002). "Regulatory circuitry of the CsrA/CsrB and BarA/UvrY systems of *Escherichia coli*." J Bacteriol **184**(18): 5130-5140.
Swanson, M. S. and R. R. Isberg (1995). "Association of *Legionella pneumophila* with the macrophage endoplasmic reticulum." Infect Immun **63**(9): 3609-3620.
Swanson, M. S. and R. R. Isberg (1996). "Identification of *Legionella pneumophila* mutants that have aberrant intracellular fates." Infect Immun **64**(7): 2585-2594.
Sze, C. W., D. R. Morado, et al. (2011). "Carbon storage regulator A (CsrA(Bb)) is a repressor of *Borrelia burgdorferi* flagellin protein FlaB." Mol Microbiol **82**(4): 851-864.
Thomason, L., D. L. Court, et al. (2007). "Recombineering: genetic engineering in bacteria using homologous recombination." Curr Protoc Mol Biol **Chapter 1**: Unit 1 16.
Thomason, L. C., N. Costantino, et al. (2007). "Multicopy plasmid modification with phage lambda Red recombineering." Plasmid **58**(2): 148-158.
Timmermans, J. and L. Van Melderen (2010). "Post-transcriptional global regulation by CsrA in bacteria." Cell Mol Life Sci **67**(17): 2897-2908.
Troutt, A. B., M. G. McHeyzer-Williams, et al. (1992). "Ligation-anchored PCR: a simple amplification technique with single-sided specificity." Proc Natl Acad Sci U S A **89**(20): 9823-9825.
Vakulskas, C. A., A. H. Potts, et al. (2015). "Regulation of Bacterial Virulence by Csr (Rsm) Systems." Microbiol Mol Biol Rev **79**(2): 193-224.
Vogel, J. P., H. L. Andrews, et al. (1998). "Conjugative transfer by the virulence system of *Legionella pneumophila*." Science **279**(5352): 873-876.
Waldor, M. K. (2010). "Mobilizable genomic islands: going mobile with *oriT* mimicry." Mol Microbiol **78**(3): 537-540.
Wang, X., A. K. Dubey, et al. (2005). "CsrA post-transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*." Mol Microbiol **56**(6): 1648-1663.

Wang, Y., N. B. Shoemaker, et al. (2004). "Regulation of a Bacteroides operon that controls excision and transfer of the conjugative transposon CTnDOT." J Bacteriol **186**(9): 2548-2557.
Wee, B. A., M. Woolfit, et al. (2013). "A distinct and divergent lineage of genomic island-associated Type IV Secretion Systems in *Legionella*." PLoS One **8**(12): e82221.
Wei, B. L., A. M. Brun-Zinkernagel, et al. (2001). "Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*." Mol Microbiol **40**(1): 245-256.
Weissenmayer, B. A., J. G. Prendergast, et al. (2011). "Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs." PLoS One **6**(3): e17570.
Wiater, L. A., A. B. Sadosky, et al. (1994). "Mutagenesis of *Legionella pneumophila* using Tn903 dlllacZ: identification of a growth-phase-regulated pigmentation gene." Mol Microbiol **11**(4): 641-653.
Wozniak, R. A. and M. K. Waldor (2010). "Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow." Nat Rev Microbiol **8**(8): 552-563.
Yakhnin, A. V., C. S. Baker, et al. (2013). "CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage." Mol Microbiol **87**(4): 851-866.
Yakhnin, A. V., J. J. Trimble, et al. (2000). "Effects of mutations in the L-tryptophan binding pocket of the Trp RNA-binding attenuation protein of *Bacillus subtilis*." J Biol Chem **275**(6): 4519-4524.
Yakhnin, H., C. S. Baker, et al. (2011). "CsrA represses translation of *sdiA*, which encodes the N-acylhomoserine-L-lactone receptor of *Escherichia coli*, by binding exclusively within the coding region of *sdiA* mRNA." J Bacteriol **193**(22): 6162-6170.
Yakhnin, H., P. Pandit, et al. (2007). "CsrA of *Bacillus subtilis* regulates translation initiation of the gene encoding the flagellin protein (*hag*) by blocking ribosome binding." Mol Microbiol **64**(6): 1605-1620.
Yang, T. Y., Y. M. Sung, et al. (2010). "Posttranscriptional repression of the *cel* gene of the ColE7 operon by the RNA-binding protein CsrA of *Escherichia coli*." Nucleic Acids Res **38**(12): 3936-3951.
Yasbin, R. E. and F. E. Young (1974). "Transduction in *Bacillus subtilis* by bacteriophage SPP1." J Virol **14**(6): 1343-1348.
Yu, D., H. M. Ellis, et al. (2000). "An efficient recombination system for chromosome engineering in *Escherichia coli*." Proc Natl Acad Sci U S A **97**(11): 5978-5983.
Yu, V. L. and J. E. Stout (2010). "*Legionella* in an ice machine may be a sentinel for drinking water contamination." Infect Control Hosp Epidemiol **31**(3): 317; author reply 318.
Zusman, T., O. Gal-Mor, et al. (2002). "Characterization of a *Legionella pneumophila relA* insertion mutant and toles of RelA and RpoS in virulence gene expression." J Bacteriol **184**(1): 67-75.