Characterization of the role of the small regulatory RNA (sRNAs) lpr0050 and lpr0024 in *Legionella pneumophila*

Malak Sadek

Natural Resource Sciences Department

McGill University
Montreal, Canada

Supervisor: Prof. Sébastien P. Faucher

June , 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Abstract

Legionella pneumophila is a facultative intracellular pathogen, and it is the causative agent of Legionnaires' disease, a life-threatening form of pneumonia. *L. pneumophila* is commonly found in most water systems including freshwater bodies, rivers, and lakes as well as in engineered water systems and man-made water distribution systems such as cooling towers. Free-living amoeba in aquatic environments are the primary reservoir of *L. pneumophila*. Following inhalation of contaminated water droplets by humans, *L. pneumophila* infects and replicates within lung alveolar macrophages and potentially causes Legionnaires' disease (LD) in susceptible individuals. The bacterium establishes its intracellular niche by forming the *Legionella*-containing vacuoles (LCVs). *L. pneumophila* governs the formation of the LCV and intracellular growth through the Icm/Dot type IVB secretion system. Icm/Dot is able to translocate around 300 protein effectors in the host cell allowing *L. pneumophila* to modulate many signalling and metabolic pathways of the host to its benefit. It is believed that Small Regulatory RNAs (sRNAs) are major players of regulation of virulence-related genes in *L. pneumophila.* We investigated the regulatory role of the two sRNA Lpr0050 and Lpr0024. Lpr0050 is encoded on the complementary strand of the effector SdeA. Using northern blot we showed that the cis-encoded sRNALpr0050 is expressed in the Exponential (E) phase and Post-Exponential (PE) phase in the wild-type, *ΔcpxR* and *ΔletS*. We showed by qPCR that the SdeA effector and its homologs (SdeB, SdeC and SidE) are co-regulated by Lpr0050. On the other hand, our results showed that the trans-encoded sRNA Lpr0024 is regulated by LetA/S in the PE phase. Furthermore, the analysis of the genomic region of Lpr0024 shows that it is present in multiple copies and uncovers a new mobile genetic element and a new endonuclease.

Résumé

Legionella pneumophila est un pathogène intracellulaire facultatif et la cause de la maladie du Legionnaire, une forme mortelle de pneumonie. *L. pneumophila* est retrouvé couramment dans la plupart des systèmes hydriques tel que les rivières et les lacs, ainsi que dans les systèmes fabriqués par l'homme comme les tours de refroidissement. Les amibes qui vivent dans ces environnements sont le réservoir primaire de *L. pneumophila*. Suite à l'inhalation de gouttelettes d'eau contaminé par l'humain, la bactérie entre dans les poumons et infectent les macrophages alvéolaires. La bactérie établit sa niche intracellulaire en format la « *Legionella*-containing vacuole » (LCV). *L. pneumophila* contrôle le développement de la LCV et sa croissance intracellulaire grâce au système de sécrétion de type IVb appelé Icm/Dot. Ce système est capable de transloquer environ 300 protéines, appelé effecteur, qui modulent plusieurs systèmes de signalement et des voies métaboliques au bénéfice de *L. pneumophila*. Les petits ARN non-codants (sRNA) ont des rôles importants dans la régulation de gènes de virulence chez *L. pneumophila*. Nous avons investigué le rôle des deux sRNAs Lpr0050 et Lpr0024. Lpr0050 est encodé sur le brin complémentaire de l'effecteur SdeA. Lpr0050 est exprimé en phase exponentielle (E) et en phase post-exponentielle (PE). Nous avons montré que SdeA et ses homologues (SdeB, SdeC et SidE) sont co-régulés par Lpr0050. D'un autre côté, l'analyse de la région génomique autour de Lpr0024 a permis d'identifier un nouvel élément génétique mobile ainsi qu'une nouvelle endonucléase.

Acknowledgments

This work would not have been possible without the academic and financial support that have been provided by my thesis advisor, Dr. Faucher. I would like to express my gratitude Dr. Faucher for giving me the chance to join his lab as a summer student in 2017 and for giving me the chance to continue as a master student. His tremendous support, guidance and scientific expertise greatly helped me in all aspects of my master's research and writing this thesis.

Nobody has been more supportive during the course of this project than the members of my family. I would like to thank my parents; whom love and guidance are with me in whatever I pursue. My parents, no words could describe all you have done for me and how I'm grateful to you! Maybe you didn't give me everything, but with no doubt you gave me all you could.

I'm thankful to Nilmimi Mendis (PhD) for her great support and patience. Nilmini was the first one to train me in the Dr. Faucher lab, she was the first one teach me basic research skills. She was always there whenever I needed her help even after she left the lab. Nilmini, you are a great researcher and an amazing person.

Durai, I cannot miss the opportunity to thank you for your help and your advice. I cannot remember even one time you refused to answer any of my questions or even to listen to me whenever I felt down. Thank you for your encouragement and for the positive vibes that you were always sending to me.

I would love to thank all the lab members of the Faucher lab for their help and support, it was a pleasure working with you and being part of your team.

I would like also to take this chance and thank Dr. Brian Driscoll. Thank you so much for having your door always opened for me and for accommodate me since my first day on MacDonald campus. Thanks for your advices and encouragement over the years. I would like to thank you for your insightful comments and feedback that helped me to develop critical thinking skills.

Julie Major, I cannot miss to thank you for being part of my academic success as you have been always a great academic advisor. Thank you for being always there whenever I need your advice. Dr. Adel and Sherry, I will never forgot what you have done for me, I will never forget your encouragement and your help. I'm in debt to you and the words cannot describe how thankful I'm to you. Thanks Sherry for believing in me and for pushing me to be where I'm now. I cannot forget your last message to me that I will have a great future. God bless your soul!

I would like to thank every and each one of my dear friends in and out of school for your support and for taking care of me when I was down. A big thanks to all of you!

Finally, thanks to CIHR for the financial support throughout my master program.

It is with an honor I dedicate this thesis to the soul of my dad who passed away in the first semester of my masters. I would love to have him alive at this moment because I'm pretty sure that he would be happy and proud of me as usual. I tried my best to make him proud even if he is not alive anymore, but I know that he is watching over me in pride. I'm dedicating this thesis to my mom as well and would like to thank her for all of the support she gave me.

Contribution of Authors

All the work and experiments design were done under the supervision of Dr. Faucher. Malak Sadek performed all experiments and wrote the thesis. The thesis was reviewed and corrected by Dr. Faucher.

Table of content

List of abbreviations

Arf1	ADP ribosylation Factor
AYE	ACES Buffer Yeast Extract
bp	Base Pair
CDC	Center of Disease Control
CYE	Charcoal Yeast Extract
E-Phase	Exponential phase
GTP	Guanosine-5'-triphosphate
HK	Histidine Kinase
HPT	Histidine Phospho transfer
IM	Inner Membrane
LCV	Legionella-containing vacuoles
LD	Legionnaires' disease
LPS	Lipopolysaccharide
MGE	Mobile Genetic Elements
MOMP	Major Outer Membrane Protein
OD_{600}	Optical Density at a wavelength of 600 nanometers
OM	Outer Membrane
OMV	Outer Membrane Vesicles
ORF	Open Reading Frame
PE-Phase	Post-Exponential PE phase
RBS	Ribosomal Binding Site
RP	Replicative Phase

RP	Response Regulator
SDS	sodium dodecyl sulfate
sRNAs	Small Regulatory RNAs
T2SS	type II secretion system
TCS	Two-component systems
TP	Transmissive Phase
WT	Wild Type

List of Figures

List of tables

Introduction

Legionella pneumophila is a facultative intracellular pathogen and it is the causative agent of Legionnaires disease (an acute form of pneumonia) and Pontiac fever (milder flu-like illness). In Europe and North America, *L. pneumophila* is responsible for over 90% of reported Legionellosis cases. *L.pneumophila* is commonly found in almost all natural systems including freshwater bodies, rivers, and lakes as well as in the engineered water systems such as cooling towers. It is able to replicate inside amoeba in the water environment, and inside lung macrophages in humans. The bacterium establishes its intracellular niche by forming the *Legionella*-containing vacuoles (LCVs). In both macrophage and protozoan host cell, a series of events that disrupt normal endocytic trafficking are needed for the intracellular multiplication of *L. pneumophila* starting with preventing phagolysosome fusion and the acidification of the LCV. *L. pneumophila* governs the formation of the LCV and intracellular growth through the Icm/Dot type IVB secretion system which is able to translocate around 300 effectors in the host cell allowing *L. pneumophila* to modulate many signalling and metabolic pathways of the host to its benefit. *L. pneumophila* replicates inside LCV until nutrients become limited, which limitation then leads to the transition to transmissive phase where many virulence-associated traits are expressed allowing the release and transmission to new host cells. This biphasic life cycle of *L. pneumophila* is reflected by a major shift in gene expression from replicative to transmissive phase. Moreover, efficient regulation of gene expression is critical for responding to the different environmental conditions that *L. pneumophila* encounters as well as for its intracellular multiplication in host cells.

The discovery of a class of regulatory elements, called small noncoding RNAs (sRNAs) uncovered a high complexity of posttranscriptional gene regulation in both prokaryotes and eukaryotes. Small RNAs are short RNA molecules that are approximately 80 to 500 nucleotides in length. sRNAs

have regulatory role in processes, such as stress response, virulence, and competence. sRNAs are classified into cis or trans encoded sRNAs that modulate gene expression through complementarity to their adjacent or distant mRNA targets, respectively. The main objective of my research project is to characterize two of these sRNAs, the cis-encoded sRNA Lpr0050 and the trans-encoded sRNA Lpr0024. Lpr0050 is encoded complimentary to Icm/Dot substrate SdeA, therefore the first aim of this project is to test the effect of the sRNA on the expression of *sdeA* and its homologs by qPCR. Lpr0024 is a trans-encoded sRNA and its target is therefore unclear. However, it is found in multiple copies in the genome. Thus the second aim of this project is to use PCR and Sanger sequencing to test if there is recombination between the repeated copies and if that recombination leads to the excision of that region in an episomal form.

1. Literature review

1.1 *Legionella pneumophila:* an environmental and human pathogen

Legionella pneumophila is an opportunistic and facultative intracellular pathogen that causes Legionellosis disease which combines two clinically distinct forms: Legionnaires' disease (LD), an acute form of pneumonia and Pontiac fever, a milder flu-like illness (Fraser et al. 1977, Glick et al.1978). In 1977, *Legionella* spp. were first identified during a major outbreak of pneumonia during the American Legion annual three-day convention at the Bellevue-Stratford Hotel in Philadelphia (McDade et al 1977). As a result, there were 182 cases; 29 deaths and the hospitalization of 147 people (Fraser et al 1977). After extensive research and investigation, the causative agent was identified as a Gram negative bacteria that was subsequently named *Legionella pneumophila* (Brenner et al 1979). *L. pneumophila* is an aerobic Gram negative, motile and rod-shaped bacterium of the γ-proteobacterial lineage, belonging to the family Legionellaceae, genus *Legionella* (Washington et al 1996, Chien et al 2004).

L. pneumophila is characterized as a water-borne pathogen as it is found in the aquatic environments inhabiting most natural aquatic systems where it is found free-living, within biofilms, and as an intracellular parasite of aquatic amoebae (Newton et al 2010, Buchrieser and Hilbi 2019). Therefore, unlike other environmental bacteria, *L. pneumophila* can grow in extracellular niches and intracellularly (Buchrieser and Hilbi 2019). Furthermore, *L. pneumophila* has been recovered from a wide range of engineered water systems such as cooling towers, fountains, and spa baths (Morris et al 1979, Sethi et al 1983, Spitalny et al 1984). On the other hand, *L. pneumophila* is characterized as an intracellular human pathogen as they can enter the human lung and infect alveolar macrophages (Newton et al 2010, Faucher and Shuman 2011). Transmission to human usually occurs only upon inhalation of contaminated aerosols, often from exposure to showers or whirlpool baths (Newton et al 2010, Spitalny et al 1984). Recently the first

potential transmission of a highly virulent *L. pneumophila* strain from person-to-person was reported (Correia et al., 2016). To date, outbreaks of legionellosis have not been associated with natural freshwater lakes (Lau and Ashbolt 2009).

1.2 Epidemiology of Legionnaires' disease

Legionnaires' disease usually begins with a mild cough, malaise, muscle aches, low fever and gastrointestinal symptoms. At later point of the infection, high fever, alveolitis and bronchiolitis start to appear (Steinert et al 2002). Legionnaires' disease is a major reportable illness in Europe, USA, Canada, New Zealand, Japan, Singapore, and Australia (Phin et al 2014). Although to date more than 60 *Legionella* species have been identified (Buchrieser and Hilbi 2019), *L. pneumophila* is responsible for over 90% of cases, which makes it the most clinically relevant and thus, well-studied among all other *Legionella* spp. (Lau and Ashbolt 2009). *Legionella longbeachae* comes to the second place since it causes more than 30% of reported cases in Australia and New Zealand (Newton et al 2010, Swart et al 2018). Although there is a at least 16 serogroups of *L. pneumophila* identified (Benson et al 1998), *L. pneumophila* serogroup 1 is responsible for more than 84% of Legionnaires' cases worldwide (Marston et al 1997, Yu et al 2002).

The case-fatality rate of Legionellosis ranges between 10 and 40% and can reach up to 50% in nosocomial outbreaks since persons at risk are usually elderly and immuno-compromised people (Buchrieser and Hilbi 2019, Hilbi 2014). Epidemiological studies indicate that there is approximately 8000 to 18000 patients hospitalized with Legionnaires' disease in the United States each year (Bartley et al 2016). The average length of stay in hospitals is estimated to be 10.2 days and average cost of $26 741 to $38 363 was estimated (Collier et al 2013). Moreover, data from Center of Disease Control and Prevention (CDC) surveillance and reporting systems indicate that the incidence of legionellosis has been increasing (Dooling et al 2011, Beer et al 2015), with a significant increase around 3.5-fold increase between 2000 and 2011 in the United States (Dooling

et al 2011). Additionally, during the period between 2000–2014, CDC investigated 27 Legionnaires' disease outbreaks, with a total of 415 cases and 65 death (Herwaldt et Marra 2018). Moreover, a systematic review was performed to provide a summary of LD outbreaks between 2006 and 2017. The records covered 136 outbreaks with 3642 total confirmed cases and 251 total deaths (Hamilton et al 2018). The global incidence of LD is unknown, it is under-reported in many countries, mainly due to lack of awareness and poor diagnostics and surveillance systems (Chaudhry et al 2018) thus, it continues to cause disease outbreaks of public health significance (Phin et al 2014).

1.3 Environmental niches of *L. pneumophila*

L. pneumophila is an environmental bacterium ubiquitous within freshwater environments (Newton et al 2010, Hilbi et al 2011). Environmental bacteria, including *L. pneumophila*, are usually exposed to predators such as free-living protozoa and nematodes (Hilbi et al 2011). Furthermore, bacteria living in such environments are continuously competing for nutrients to be able to survive in these harsh environments (Lau and Ashbolt 2009). In these environment *L. pneumophila* can acquire nutrients by residing in relatively nutrient-rich communities called biofilms (Lau and Ashbolt 2009). Biofilm formation is favored at high temperatures (37–42°C) (Lau and Ashbolt 2009). These biofilms represent a network of metabolic interactions, which provides nutrients and protects the resident bacteria from physical, chemical and biological hazards (Hilbi et al 2011). It has been shown that within biofilms *L. pneumophila* can grow by 'necrotrophy' on dead bacteria and potentially other organic materials (Hilbi et al 2011). Interestingly, model biofilm systems have demonstrated that the replication of *L. pneumophila* within biofilm depends on the presence of a protozoan host (Murga et al 2001). Huws et al. (2005) have shown that certain protozoa are able to graze on biofilm and therefore play a critical role in the development of the biofilm (Huws et al 2005). It has been shown that *L. pneumophila* survive

their encounter with free-living protozoa, that graze on the biofilm, and establish a parasitic relationship with these hosts (Greub and Raoult 2004, Lau and Ashbolt 2009). Therefore, it is believed that free-living protozoa may serve as the main source for the spread of *L. pneumophila* by providing an intracellular environment for multiplication of these pathogenic strains (Lau and Ashbolt 2009).

1.4 Intracellular replication of *L. pneumophila*

In order to understand how *L. pneumophila* is able to cause the disease, it is important to understand the interaction of *L. pneumophila* with their host cells. This relationship has been studied for a wide variety of protozoan and mammalian host cells (Newton et al 2010). Evolutionarily, the interactions between *L. pneumophila* and their protozoan hosts have led to adaptive responses, which in turn makes the bacterium able to also infect mammalian phagocytes such as human alveolar macrophages (Swart et al 2018). Amoebae-resistant *L. pneumophila* are also resistant to macrophages, which is a key for the bacteria to cause disease in humans (Hilbi et al 2011).

It is known that free-living protozoa share many characteristics with the mammalian phagocytes, in particular macrophages (Hilibi 2011). Given the similarity of the infection process in amoebae and macrophages, the amoebae have been a strong model organism to study *L. pneumophila*-macrophage interactions (Hilbi et al 2007, Escoll et al 2013, Hoffmann et al 2013). In particular, *Acanthamoeba castellanii*, *Hartmannella vermiformis*, *Naegleria spp*., and *Dictyostelium discoideum* have been widely used to study the relationship of *L. pneumophila* with their environmental hosts (Steinert et al 2005). For mammalian cells, much work has been performed using macrophage-like cultured cells and murine bone marrow-derived macrophages (Newton et al 2010).

When *L. pneumophila* is engulfed by free-living amoeba or by lung alveolar macrophages, it replicates intracellularly within a well-developed vacuole called *Legionella*-containing vacuole (LCV) (Escoll et al 2013). Within this protected vacuole, *L. pneumophila* replicates and becomes acid tolerant (Jules and Buchrieser 2007). Successful intracellular multiplication of *L. pneumophila* requires a set of well-ordered and organized events that modulate the normal endocytic trafficking pathways in both macrophages and protozoan host cells (Hovel-Miner et al 2009). These events include preventing the acidification of LCV and fusion with lysosomes, thus preventing the maturation of the LCV into a phagolysosome (Horwitz and Maxfield 1984).

Moreover, the establishment of the LCV requires remodeling of the LCV by recruiting the early secretory vesicles, mitochondria, and vesicles derived from the Golgi and from the endoplasmic reticulum (ER) (Isberg et al 2009, Escoll et al 2013, Faucher and Shuman 2011). The early secretory vesicles that transport between the ER and the Golgi are disrupted, their luminal content are merged into the LCV, an ER-like organelle is created to support replication of *L. pneumophila* (Derré and Isberg 2004, Kagan and Roy 2002). Interestingly, when intact cells and isolated LCVs were analyzed, ER-associated proteins were found located close to the LCV shortly after *L. pneumophila* uptake (Derré and Isberg 2004, Kagan et al 2004). Among these proteins is Sec22b, which is a member of the membrane fusion proteins family SNARE (Derre et al 2004). Furthermore, Rab1 is a small GTPase that regulates trafficking from ER to Golgi (Isberg et al 2009).

It has been reported that one hour after infection, the LCV is surrounded by smooth vesicles and by at least one mitochondrion close to the vacuolar membrane (Escoll et al 2013). Mitochondria recruitment around the LCV has also been shown after *L. pneumophila* infection of the amoeba *D. discoideum*, *N. fowleri* and H. *vermiformis* (Abu Kwaik 1996, Francione et al. 2009).

1.5 Type IV secretion system and Dot/Icm effectors

Of the most important virulence factors of *L. pneumophila* is the Type IV secretion systems (T4SSs) which are evolutionarily related to bacterial conjugation system (Escoll et al 2013, Halbi et al 2011). The most important T4SS for *L. pneumophila* is the Dot/Icm T4bSS. The doubled name was because the Isberg and Shuman groups independently discovered it and named it Dot/Icm (defective in organelle trafficking/intracellular multiplication system) (Marra et al 1992, Berger et al 1994). Shuman group discovered the locus that was able to restore the wild type phenotype of the mutant with defective intracellular multiplication and named it *icm* (Marra et al 1992). On the other hand, Isberg group discovered the locus that was able to restore the wild type phenotype of the mutant with a defect in organelle trafficking calling it *dot* (Berger et al 1994).

The Dot/Icm system is made up of approximately 27 proteins (Trigui et al 2013, Hovel-Miner et al 2009). It is essential for intracellular replication of *L. pneumophila* in both, amoeba and macrophages (Wood et al 2015, Segal and Shuman 1999). Moreover, LCV formation and other *Legionella*–host interactions are mainly controlled by Dot/Icm (Kubori and Nagai 2016 , Grohmann et al 2018, Isberg et al 2009). It has been reported that mutations in many of these genes cause defective recruitment of ER-derived material to the LCV (Isberg et al 2009).

The major function of the Dot/Icm system is to deliver proteins, known as effectors, across the host cell membrane into the host cytosol (Isberg et al 2009, Backert and Grohmann 2018). These effectors function to remodel LCV and make it a safe environment for replication (Wood et al 2015). Dot/Icm translocates nearly 300 different effectors into host cells (Hubber and Roy 2010, Finsel and Hilbi 2015). About 10 % of *L. pneumophila* genome code for these translocated effectors (Al-Quadan et al 2012). Moreover, whole genome sequencing has revealed that the presence of these effectors is quite variable among the sequenced genomes.

Nagai et al (2002) first reported that the effector protein RalF is secreted and translocated by the Dot/Icm system (Nagai et al 2002). Following this finding, several Dot/Icm substrates were discovered via multiple genetic screens and bioinformatic approaches (Campodonico et al 2015, Luo et al 2004, Nagai et al 2002).

Dot/Icm substrates often carry conserved protein domains mostly found in eukaryotic organisms, such as serine–threonine kinase, ubiquitin ligase, Sel-1, Sec7, U-box, F-box, ankyrin repeats (Cazalet et al 2004; de Felipe et al 2005). This suggests that *L. pneumophila* acquires foreign genetic material from eukaryotic cells that is incorporated into its genome and therefore, retains part of its original activity (de Felipe et al 2005). These proteins are likely acquired by horizontal gene transfer (HGT), indicating that some effector proteins may have been acquired through horizontal gene transfer (de Felipe et al 2005). de Felipe et al performed a bioinformatic screen for eukaryotic motifs in all open reading frames of the *L. pneumophila* Philadelphia-1 genome. They identified 44 uncharacterized genes and confirmed that seven of these gene products were translocated into host cells through Dot/Icm (de Felipe et al 2005).

Screening for mutant defective for intracellular growth was used as approaches to identify Dot/Icm substrates (Isberg et al 2009). This method has only identified a few genes, the best example is *sdhA* whose deletion generally blocks intracellular growth (Laguna et al 2006). Additionally, there are other methods used to identify Dot/Icm substrates such as the use of bioinformatics analysis to identify proteins that have similar sequences to those involved in the eukaryotic processes (Isberg et al 2009), these *Legionella* -like- eukaryotic genes are called *leg* and they include kinases, lyases and esterases (Cazalet et al 2004). Moreover, identification of *L. pneumophila* proteins that disrupt cellular processes in the yeast *Saccharomyces cerevisiae can be used* to identify Dot/Icm

substrates (Campodonico et al 2005). The identification of regulatory networks that control the translocated effectors is another approach of Dot/Icm substrates identification (Zusman et al 2007). The main function of these translocated effectors is to control different steps in the infection by targeting and modulating many cellular pathways and different eukaryotic organelles, including endocytosis, gene expression, protein production, transcription factors (Escoll et al 2016, Qiu and Luo 2017). Additionally, many effectors contribute to evade phagosome maturation and to promote fusion of the LCV with the ER-derived vesicles (Wood et al 2015). Furthermore, *L. pneumophila* actively modulates the vesicle trafficking pathway through the activity of Dot/Icm effectors (Qiu and Luo 2017).

A key to disrupt trafficking events within eukaryotic cells is the ability of *L. pneumophila* to recruit GTPases and control GTP cycling (Newton et al 2010). Dot/Icm effectors act on different organelles of the infected cell to manipulate small GTPases and other molecules that are important for vesicle trafficking (Newton et al 2010). For example, the Dot/Icm substrate LegG1 activate the small GTPase Ran (Rothmeier et al 2013). Arf1 (ADP-ribosylation factor) and Rab1 are small GTPases whose requirement is important for vesicle transport between the ER and the Golgi apparatus (Hardiman et al 2012) therefore, it was hypothesized that they are potential targets of the Dot/Icm substrates. The translocated substrate RalF activates *arf1, ralF* mutant was defective for recruitment of *arf1* to the LCV (Isberg et al 2009). Rab1, the GTPase that promotes the fusion of ER-derived vesicles with Golgi compartments, is recruited to the LCV within minutes of bacterial uptake (Derre et al 2004, Kagan et al 2004). The recruitment of Rab1 to the LCV is regulated by the effector SidM (Qiu and Luo 2017). The screening for mutants defective in Rab1 recruitment identified several other effector such as LepB (Murata et al 2006).

The Dot/Icm T4SS is not only required for the establishment of the LCV, but also involved in other processes in the infection including bacterial entry, the inhibition of host cell apoptosis, autophagy, cell death pathways and the egress of *L. pneumophila* from host cells (Hilbi et al 2001).

1.5.1 SidE effector proteins

The SidE effector family contains four large effector proteins SidE, SdeA, SdeB, and SdeC that are translocated into macrophages via the Dot/Icm system (Isberg et al 2004, Bardill et al 2005). Three members of this family, *sdeA, sdeB, and sdeC*, are organized in an operon-like structure, where they are located on the same locus *sdeC-sdeA* (Liu et Luo 2007). *sidJ* (*lpg2155*) and *lpg2154* are located between the *sdeC* and *sdeB* genes but they are not part of SidE family, while the fourth member *sidE* is located in a different area on the chromosome (Luo and Isberg 2004, Jeong et al 2015). The *sidE* family is predicted to encode proteins of approximately 170 kDa in size, with the exception of *sdeB*, which is a little larger and is predicted to encode a 200 kDa protein.

1.5.1.1 Function of SidE family effectors

SidE family effectors are required for virulence and intracellular growth of *L. pneumophila* (Bardill et al 2005). A strain with a complete *sidE* family deletion was unable to grow to wild-type levels exhibiting approximately 10-fold less growth than wild-type (Bardill et al 2005, Jeong et al 2015). In 2007, Liu and Luo studied the role of the three genes *sdeA*, *sdeB*, and *sdeC* in *L. pneumophila* intracellular replication (Liu and Luo 2007). They constructed a strain with a deletion of the entire region from *sdeC* to *sdeA*; this mutant exhibited about 20-fold decrease in the intracellular replication, intracellular defect could not be complemented by either *sdeA*, *sdeB*, or *sdeC* alone (Liu et Luo 2007). However, *sidJ* was used to complement the mutant and the growth defect was partially restored, indicating that this gene is responsible for the observed growth defect (Liu et Luo 2007). Jeong et al (2015) showed that expression of SidJ and at least one member of the SidE family is required for a full virulence of *L. pneumophila* within *A. castellanii* (Jeong et al

2015). Overall, these results suggest an important role of the SidE effectors, however, their actual function remains unclear because the lack of homology of the SidE family with proteins of known function makes it difficult to predict their function (Bardill et al 2005).

1.5.1.2 SidJ is a metaeffector

The characterization of SidJ provides an example of metaeffector or effector–effector interactions (Hilbi et al 2017). Metaeffector is a class of effectors which targets another effector protein instead of host factors (Hilbi et al 2017). Given that SidJ and SdeC, SdeB, and SdeA are all expressed from the same locus, it is believed that there is a functional connection between these proteins (Jeong et al 2015). In an extensive study, Jeong and co-workers proposed that the molecular connection between SidJ and SidE family members occurs within host cells as they found that SidJ does not seem to regulate the secretion of SidE family members and that SidE proteins do not influence the export of SidJ (Jeong et al 2015).

SidE proteins are translocated into the host cells by the Dot/Icm at early points during the infection and localized on the cytoplasmic face of LCV (Bardill et al 2005). As the infection proceeds, SidJ begins to accumulate in the host cell, eventually reaching a critical peak (Liu et Luo 2007). There is correlation between the gradual accumulation of SidJ during the infection and the decreased level of the SidE family on the LCV (Jeong et al 2015). This observation suggests that SidJ directly promote the removal of SidE proteins from the LCV. Additionally, Qui et al (2017) suggested that SidJ may modulate the function of SidE when they found that the expression of SidE proteins is toxic in both yeast and mammalian HEK293 cells, but this toxicity can be repressed via the co-expression of SidJ (Qui et al 2017). In summary, the Dot/Icm substrate SidJ functions as a metaeffector which regulates the activity of other effector protein such as SidE (Kubori et al 2010).

1.5.1.3 SidE are involved in host ubiquitination

SidE family play major role in modulating host ubiquitination pathways (Sheedlo et al 2015, Qui et al 2016, Qui et al 2017). Ubiquitin, a small 76-aa protein modifier, is involved in most of the eukaryotic cellular processes (Sheedlo et al 2015). Ubiquitination, the covalent attachment of ubiquitin to a substrate protein, is catalyzed by the E1, E2 and E3 three-enzyme cascade which links the carboxy terminus of ubiquitin to the ε-amino group of the substrate through an isopeptide bond (Qui et al 2016). The three-enzyme cascade includes deubiquitinases (DUBs) whose function is to remove ubiquitin from the modified substrates (Hershko 2000). Ubiquitination is a post-translational modification and is considered one of the most essential signaling mechanisms in eukaryotes since it is involved in many cellular processes such as cell cycle, development and immunity (Popovic et al 2014). Thus, disrupting the host ubiquitin signaling is critical for the success of many infectious agents. Many bacterial pathogens use virulence factors to interfere with the host ubiquitin pathway for a successful infection (Zhou et al 2015).

Members of the SidE effector family of *L. pneumophila* ubiquitinated multiple Rab small GTPases associated with the endoplasmic reticulum (Luo et al 1999). SdeA induces a biochemical modification of multiple ER-associated Rabs, and at least Rab33b and Rab1 are substrates during bacterial infection without the need of the E1 and E2 enzymes (Qui et al 2016). Therefore, the study of the SidE effectors of *L. pneumophila* has uncovered a new mechanism of ubiquitination that does not require the E1 and E2 enzymes (Qui et al 2017).

On the other hand, extensive study by Qui et al (2017) showed that SidJ reverses ubiquitination of SidE-modified substrates by cleaving the phosphodiester bond that links phosphoribosylated ubiquitin to protein substrates (Qui et al 2017). Therefore, SidJ has deubiquitinase activity against ubiquitination induced by the conventional three-enzyme cascade. This deubiquitinase activity of SidJ is essential for its role in *L. pneumophila* infection. Finally, the activity of SidJ is required to

efficiently reduce the flood of ubiquitinated Rab33b in the infected cells within a few hours after bacterial uptake. These results propose that SidJ is a ubiquitin-deconjugating enzyme that functions to regulate the activity of SidE effectors (Qui et al 2017).

1.6 Other Virulence factors

L. pneumophila possesses several virulence factors that are involved in every step in the infection cycle (Zhan et al 2015). Theses bacterial determinants that are essential for *L. pneumophila* pathogenicity include lipopolysaccharide (LPS), flagella, pili, a type II secretion system (T2SS), and outer membrane proteins (Newton et al 2010).

1.6.1 Attachment and Entry

The attachment and entry of *L. pneumophila* to the host cell surface is the first and most important step in the infection cycle. (Escoll et al 2013, Zhan et al 2015). Several proteins and bacterial factors are involved in the entry of *L. pneumophila* into host cells including PilEL, RtxA, EnhC, and LpnE (Cirillo et al 2000, Cirillo et al 2001, Escoll et al 2013). The *pilEL* gene, responsible for the expression of a long pili (discussed below) in *L. pneumophila*, plays important role in the attachment of *L. pneumophila* to *Acanthamoeba polyphaga* (Stone and Abu Kwaik 1998). RtxA is involved in the attachment and entry of *L. pneumophila* into *Acanthamoeba castellanii* and it mediates *L. pneum*ophila attachment to human cells, since *rtxA* mutants exhibited a decreased adherence and entry into human epithelial and monocytic cell lines (Cirillo et al 2000). EnhC is a periplasmic protein and it is important in the adherence of *L. pneumophila* to *A. castellanii* as well as for attachment and entry into human epithelial and monocytic cell lines (Cirillo et al. 2000). Additionally, EnhC was found to be essential for efficient replication in macrophages and for the maintenance of cell wall integrity (Liu et al 2012). LpnE is required for full entry of *L. pneumophila* into the macrophage cell line THP-1, and the epithelial cell line A549 (Zhan et al 2015). Furthermore, LpnE is required for an efficient infection of *A. castellanii* and for the

replication of *L. pneumophila* in the lungs of A/J mice (Newton et al 2007). Other *L. pneumophila* proteins contributing to the adhesion to macrophages have been identified which include the major outer membrane protein (MOMP), collagen like protein named Lcl, and a putative *L. pneumophila* -specific adenylate cyclase named LadC (Bellinger-Kawahara and Horwitz 1990, Krinos et al 1999, Newton et al 2008).

1.6.2. Pili

One of the main determinants for attachment and entry of *L. pneumophila* into epithelial cells is a type IV pili (Zhan et al 2015, Stone and Abu Kwaik 1998). Pili have been shown on the surface of *Legionella* by transmission electron microscopy (Rodgers et al 1980). *L. pneumophila* expresses pili of variable lengths, either long (0.8 to 1.5 μm) or short (0.1 to 0.6 μm) (Stone and Abu Kwaik 1998). The PilE protein is the main constituent of long type IV pili. The encoding genes of *Legionella* pili include *pilB*, *pilC*, *pilD*, *pilE*, *pilM*, *pilN*, *pilO*, *pilP* and *pilQ* (Zhan et al 2015). Strains lacking the type IV pilin gene showed disappearance of long pili on the surface of *L. pneumophila* and nearly a 50% decrease in the adherence to human epithelial cells, macrophages and *A. polyphaga* (Stone and Abu Kwaik 1998). Additionally, this pili is involved in the formation of biofilm (Coil and Anné 2009).

1.6.3 Flagella

L. pneumophila possesses a single monopolar flagellum (Liles et al., 1998; Stone and Abu Kwaik, 1998). Rodgers and co-workers used negative stain electron microscopy to examine strains of *L. pneumophila*, and they provided the first report that these strains possess both flagella and pili (Rodgers et al 1980).This organelle plays an important role in cell motility, adhesion, and host invasion and it is also involved in biofilm formation (Heuner and Albert-Weissenberger 2008). Importantly, switching from the replicative phase to the transmissive phase (section 1.7) is co-regulated with the expression of flagella (Jules and Buchrieser 2007, Molofsky et al 2004).

Regulators that control flagellation also regulate other important virulence traits such as lysosome avoidance and cytotoxicity (Shevchuk et al 2011).

1.6.4 Cell envelope

The bacterial envelope plays important roles in the pathogenesis of *L. pneumophila* (Zhan et al 2015). It mainly protects the bacterium from environmental hazards, allows a passage of nutrients and waste products in and out of the cells, and mediates the direct contact with other organisms (Shevchuk et al 2011). The inner membrane (IM) and outer membrane (OM), periplasm, Lipopolysaccharides (LPS), outer membrane vesicles (OMV) are components of *L. pneumophila* cell envelope and they all mediate its virulence (Shevchuk et al 2011, Cianciotto 2001).

An important function of the inner membrane of *L. pneumophila* is the regulation of iron acquisition (Cianciotto 2007). Iron uptake is an essential process during all phases of *L. pneumophila* growth (Shevchuk et al 2011). Furthermore, the inner membrane proteins often regulate different cytoplasmic processes such as gene expression (Shevchuk et al 2011).

The outer membrane (OM) is a special characteristic of Gram-negative bacteria; it is a lipid bilayer composed of phospholipids, lipoproteins, LPS, and proteins (Shevchuk et al 2011). The outer membrane is where mature LPS molecules and the shedding of OMVs are located (Zhan et al 2015, Cianciotto 2001). Strains lacking phosphatidylcholine from the *L. pneumophila* envelope causes reduced cytotoxicity and multiplication of the bacterium inside macrophages (Conover et al 2008). Furthermore, the strains lacking this lipid do not efficiently bind to macrophages (Palusinska-Szysz et al 2011).

Some of the outer membrane protein components are directly involved in adhesion and invasion (Shevchuk et al 2011). The major outer membrane protein (MOMP) which is encoded by *ompS* gene, contributes to the attachment to host cells (Zhan et al 2015). Moreover, the heat shock protein Hsp60 is also important for attachment and invasion (Garduño et al, 1998). Hsp60 is a member of

the GroEL family molecular chaperones and is synthesized particularly during intracellular growth (Garduño et al, 1998). Finally, Mip, the macrophage infectivity potentiator, is a membrane-associated homodimeric protein which is found on the bacterial surface and is essential for replication within host cells efficiently (Riboldi-Tunnicliffe et al, 2001, Wagner et al, 2007).

Lipopolysaccharides (LPS) are found on the outer membrane and consist of a lipid and a polysaccharide composed of O-antigen based on which *L. pneumophila* can be divided into at least 15 serogroups (Helbig and Amemura-Maekawa, 2009). LPS are considered a virulence factor, *L. pneumophila* LPS interacts with the host cells and modulate intracellular trafficking (Shevchuk et al 2011). Moreover, LPS production is correlated with serum resistance and intracellular growth (Lüneberg et al, 1998).

Outer membrane vesicles (OMVs) are shed from the outer membrane (Zhan et al 2015). *L. pneumophila* OMVs contain a large number of virulence-associated proteins (Galka et al, 2008) OMVs adhere to cell surfaces and can translocate toxins and other virulence factors. Additionally, the immune response, cytokine profile, inflammation and innate immunity are modified by contact with OMVs (Ellis and Kuehn, 2010).

1.6.5 Type II secretion systems

In addition to the Dot/Icm T4SS, *L. pneumophila* also possesses a type II protein secretion system (T2SS) known as Lsp. (Hales and Shuman 1999, Soderberg et al 2008). Substrates of T2SSs are translocated across the inner membrane via Sec or Tat machinery (Newton et al 2010). The first evidence that *L. pneumophila* possess T2SS was when pilD, which processes the pilin of type IV pili, was discovered as one of the T2SS components that encodes the pseudo pilin peptidase (T2S O) (Liles et al, 1999). *L. pneumophila* T2SS has 12 components located in five loci throughout the chromosome (Newton et al 2010). Initially, 12 substrates were demonstrated to be secreted by the T2SS (Cianciotto 2005, Rossier et al 2004) suh as ProA and ChiA that are associated with the LCV

membrane (Truchan et al 2017). Lsp is required for full virulence and environmental persistence (Newton et al 2010). Mutant of the *lsp* system are defective for replication of *L. pneumophila* in the lungs of mice compared to wild-type indicating that Lsp plays a role in the intracellular growth (Rossier et al 2004, Zhan et al 2015). On the other hand, a disruption of the Lsp secretion system makes *L. pneumophila* unable to survive at low temperatures ranging from 4°C to 17°C in tap water, and from 22°C to 25°C in the presence of aquatic amoebae (Soderberg et al 2008). This observation suggests that Lsp plays a role in the survival of *L. pneumophila* at low-temperature (Zhan et al 2015).

Table 1. List of *L. pneumophila* virulence factors other than the TIVSS and its effectors

Nmae	Function	Reference
PilE	Expression of the long pili. attachment of *L. pneumophila* to *Acanthamoeba polyphaga*	(Stone and Abu Kwaik 1998)
RtxA	adherence and invasion of host cells	(Cirillo et al 2000).
EnhC	maintenance of cell wall integrity, facilitating *Legionella* intracellular growth	(Cirillo et al 2000).
LpnE	mediate *L. pneumophila* attachment and entry to human cells	(Zhan et al 2015)
Lcl	enhances invasion and cytokine expression, recruitment of mitochondria to the nascent LCV	(Krinos et al 1999)
LadC	adhesion to macrophages, intracellular replication, putative modification of protein functions	(Newton et al 2008)
MOMP	attachment to host cells	(Horwitz 1990)
Hsp60	Attachment to and invasion of a HeLa cell, mediate phagocytosis of *L. pneumophila*	(Garduño et al, 1998)

Mip	Efficient replication within host cells and transmigration	(Riboldi-Tunnicliffe et al, 2001)
type IV pili	entry of *L. pneumophila* into macrophage, influenced trafficking of the *L. pneumophila* vacuole	(Zhan et al 2015).
LPS	interacts with the host cells and modulate intracellular trafficking	(Shevchuk et al 2011)
Flagella	motility	(Rodgers et al 1980).
LsP	Type II secretion system, virulence and environmental persistence	(Hales and Shuman 1999)

1.7 *L. pneumophila* life cycle

In natural and drinkable water supplies, *L. pneumophila* stays within biofilm communities where it is grazed by amoebae (Lau and Ashbolt 2009). When *L. pneumophila* are ingested by protozoan hosts or macrophages, it resists digestion and replicates actively within the LCV and eventually it kills its host and returns to the aquatic environment searching for new host (Molofsky and Swanson 2004). The life cycle of *L. pneumophila* within *A. castellanii* mirrors that found in alveolar macrophages (Swanson and Hammer 2000). Once the nutrients are depleted and the LCV is no longer a suitable niche for replication, switching to the transmissive phase is required which is coordinated through a regulatory cascade leading to the expression of traits that promote transmission to a new host cell (Bruggemann et al 2006).

Therefore, *L. pneumophila* exhibits a biphasic life style, switching between a replicative phase (RP) which does not express transmission and virulence traits and a transmissive phase (TP) which is an infectious, non-replicating form (Molofsky and Swanson 2004, Byrne and Swanson, 1998). In the laboratory this can be represented by growing the bacteria in AYE broth, where the exponential phase (E) mimics RP, and the post-exponential stationary phase (PE) models TP

(Molofsky and Swanson 2003). During RP *L. pneumophila* is less virulent, sodium resistant and not flagellated; however, once *L. pneumophila* enters TP, or PE in broth cultures, the bacteria express transmission traits which include: (1) motility; (2) resistance to the stresses such as heat, osmotic pressure and nutrient limitation; and (3) virulence factors to prevent lysosome fusion (Molofsky and Swanson 2004). These traits are not expressed in the RP or E phase. After *L. pneumophila*, establishes another intracellular niche successfully, it goes back to the RP in order to start a new cycle (Molofsky and Swanson 2004).

Successful transition from replicative to transmissive phase and from the extracellular to the intracellular environments requires a precise adaptation to the different conditions that *L. pneumophila* encounters (Jules and Buchrieser 2007). The global expression profiles of *L. pneumophila* strains showed genes, around 405, that are upregulated at earlier points during the infection (RP), around 393 genes that are upregulated at late points (TP) (Jules and Buchrieser 2007).

The RP genes included DNA and RNA polymerase complexes, transcription and translation elongation factors, ribosomal proteins, protein secretion and translocation systems (secAB, secEF, secGY, etc.) and other genes encoding metabolic processes. On the other hand, among the genes induced during TP, there are several regulators (e.g. *fliA*, *cpxR*, *rpoE*), virulence factors (e.g. *ralF*, *dotA*, *letE*, *enhA*, *sdeA*, *sdcA*) and flagella biosynthesis genes (e.g. *fliS*, *fliD*, *fliN*) (Bruggemann et al 2006). Together, these results confirm that during RP bacteria is less virulent and the transcriptome during RP shows that cell division and intracellular multiplication is active. In contrast, during TP, *L. pneumophila* becomes more virulent which prepares it to evade the host cell and to search for a new one.

Amino acids (especially Ser, Thr, and Glu) are major sources of carbon, nitrogen, and energy for *L. pneumophila* (Eisenreich et al 2015). Transcriptome analysis showed that 1) high number of the genes involved in the metabolism of amino acids, lipids, carbohydrates, nucleotides, cofactors, and vitamins, 2) many genes encoding transporters for amino acid and oligopeptides, and 3) genes encoding for enzymes involved in degradation of Lys, Arg, His, Thr, Glu, and Gln, were highly induced during growth in macrophages (Faucher et al 2011).

L. pneumophila requires large amounts of iron to grow in broth and inside host cells and the bacterium has many systems to acquire sufficient amounts of iron (Cianciotto, 2007). Iron transport systems are induced during intracellular growth (Jules and Buchrieser 2007).

1.8 Regulation of gene expression

The biphasic life cycle of *L. pneumophila* is regulated by a major shift in gene expression from replicative to transmissive phase (Bruggemann et al 2006). Efficient regulation of gene expression is necessary for responding to the different environmental conditions that *L. pneumophila* encounters as well as for its intracellular multiplication in host cells (Trigui et al 2013). Intensive research indicated that many regulators are implicated in the control of virulence-related genes and proteins through well-regulated stress response pathways (Faucher et Shuman 2011, Buchrieser and Hilbi 2019, Hilbi 2014). Two-component systems (TCS) are major regulators of stress response pathways and virulence factors required during host cell infection (Faucher et Shuman 2011, Segal 2013). These include PmrA/PmrB (Zusman et al, 2007), CpxR/CpxA (Altman and Segal, 2008), LetA/LetS (Hammer et al, 2002) and LqsS/LqsR (Tiaden et al 2010). Additionally, the sigma factor RpoS (σ^S) regulates a number of known virulence factors including many Icm/Dot effectors (Hovel-Miner et al 2009) and it is required for intracellular multiplication.

1.8.1 Two component systems

Two-component systems (TCSs) are common signal transduction systems that make bacteria able to sense and respond to various environmental stimuli through changes in gene expression (Segal 2013, Gotoh et al 2010, Jung et al 2012). TCSs are widely used by many pathogenic bacteria to control the expression of their virulence genes (Segal 2013). A typical TCS is composed of a membrane-integrated sensor kinase (histidine kinase, HK), and a cytoplasmic transcriptional response regulator (RP) containing an N-terminal receiver domain and a C-terminal DNA binding domain (Segal 2013, Jung et al 2012). An environmental stimuli is detected by the sensor histidine kinase leading to its autophosphorylation. Then the phosphoryl group from the histidine residue is transferred to an aspartic acid residue in the receiver domain of the response regulator leading to its activation (Segal 2013, Gotoh et al 2010). Bacteria usually have multiple TCSs, each one responding to a specific environmental signal, such as pH, nutrient level, osmotic pressure, quorum signals and antibiotics (Gotoh et al 2010). In *L. pneumophila*, 13 histidine kinases and 14 response regulators were identified (Cazalet et al, 2004); however, to date only a few of these systems have been studied (Sahr et al 2009). Four TCSs regulate stress response pathways and virulence as well as the expression of *L. pneumophila* effector-encoding genes: CpxRA, LetAS, PmrAB and LqsRS)

1.8.1.1 The CpxRA TCS

This system consists of the CpxR cytoplasmic response regulator and the CpxA sensor histidine kinase (Raivio et al 1997). CpxA phosphorylate CpxR allowing it to bind to the DNA and modulate transcription (Segal 2013). In the absence of the suitable signals or environmental stimuli, CpxR remains dephosphorylated and therefore inactive (Segal 2013). In *E. coli*, the CpxAR system is activated by envelope stress signals: CpxA was shown to sense misfolded proteins in the bacterial envelope, leading to phosphorylation of CpxR transcriptional activation of target genes (Gal-Mor and Segal 2003).

In *L. pneumophila* , CpxRA system is a major player in virulence as it is known to regulate many virulence factors (Tanner et al 2016). Genetic screening searching for regulators of the *icmR* gene was the first evidence of the involvement of the CpxRA in *L. pneumophila* virulence and its contribution to the regulation of Dot/Icm system genes and effectors (Gal-Mor and Segal 2003). CpxR directly activate the expression of four Dot/Icm system genes: *icmR*, *icmV*, *icmW* and *lvgA* as well as the regulation of 11 effector proteins (Gal-Mor and Segal 2003, Altman and Segal 2008). An extensive study of CpxRA TCS by Tanner et al (2016) shows, via analyses of the *cpxRA* mutant strains, that CpxRA system regulates several of Dot/Icm effectors, other than the 11 effector proteins, as well as other major virulence factors including type II secreted substrates (Tanner et al 2016). They have also shown that intracellular growth in *A. castellanii* requires CpxRA (Tanner et al 2016).

1.8.1.2 LetAS TCS

LetAS TCS consists of the LetA response regulator and the LetS sensor histidine kinase (Hammer et al. 2002). In *L. pneumophila* LetAS is a major regulator of differentiation in response to starvation inside the host cells as well as in broth culture (Edwards et al 2010).The first report that LetAS TCS is involved in *L. pneumophila* virulence came from a genetic screening looking for mutants that express the flagellin gene poorly (Hammer et al. 2002). *L. pneumophila* LetAS system belongs to a family of signal-transducing proteins that is known to use multiple steps in order to regulate gene expression (Edwards et al.2010). LetS was shown to include three main domains (transmitter (T), receiver (R) and histidine phosphotransfer (HPT) domains) that are all contributing to the activation of its response regulator LetA. (Dupré et al 2015, Edwards et al. 2010). LetAS is activated by different stimuli (Edwards et al. 2010). Once LetA is activated, the T domain is phosphorylated by an ATP molecule which leads to the phosphorylation of the R

domain. Lastly, the HPT domain receives the phosphate from the R domain and sends it to LetA (Edwards et al.2010).

LetAS induces genes that enable efficient host transmission and survival in the environment, including cytotoxicity, motility, pigment production, infectivity, and lysosome evasion (Dalebroux et al 2009, Mendis et al 2018). Activated (phosphorylated) LetA positively regulates the transcription of two small regulatory RNAs, RsmY, and RsmZ, and then the transmission phase is regulated by LetAS-RsmYZ-CsrA regulatory cascade discussed below (Segal 2013, Faucher and Shuman 2011, Trigui et 2013, Sahr et al 2009). Additionally, it has been shown that LetAS is required for efficient intracellular multiplication of *L. pneumophila* (Gal-Mor and Segal 2003). Strains lacking LetA was found to be defective for intracellular growth of *L. pneumophila* in amoeba (Gal-Mor and Segal 2003, Hammer et al. 2002, Lynch et al 2003)

1.8.1.3 PmrAB TCS

The PmrAB TCS consists of the PmrA response regulator and the PmrB sensor histidine kinase (Zusman et al 2007). PmrAB was studied widely in *Salmonella enterica* where it was involved in regulating the expression of virulence genes (Gunn 2008). Moreover *S. enterica* PmrAB TCS is active when the bacteria are inside macrophages and during infection of mice (Merighi et al 2005). PmrAB respond to several environmental stimuli such as low concentrations of Mg^{+2}, high levels of Fe^{+3} and low pH (Zusman et al 2007).

The contribution of PmrAB TCS in *L. pneumophila* virulence was first identified via a bioinformatic analysis of several *L. pneumophila* effector-encoding genes with a conserved regulatory element consisting of a tandem repeat sequence (cTTAATatT) (Zusman et al 2007). Further researches have reported that a very similar sequence to that tandem repeat sequence has been previously identified in *S. enterica*, and were recognized by PmrA (Marchal et al 2004). The

environmental stimulus that activates the *L. pneumophila* PmrB is not known, but its activation is related to the pH levels of the LCV (Segal 2013). PrmAB control the Icm/Dot type IV secretion pathway (Gunn 2008). Moreover, the gene encoding for PmrA is required for intracellular growth of *L. pneumophila* in amoeba (Zusman et al 2007).

1.8.1.4 LqsRS TCS

LqsRS TCS consists of the LqsR response regulator and the LqsS sensor histidine kinase (Spirig et al 2008). The *L. pneumophila* LqsSR is homologous to the *Vibrio cholerae* quorum sensing system CqsAS (Miller et al 2002). The first identification of LqsSR TCS in *L. pneumophila* was throughout a bioinformatic analysis of the *L. pneumophila* genome to find a homologous system to the CqsAS system from *V. cholerae*; the corresponding *L. pneumophila* proteins were termed LqsA and LqsS (for Legionella quorum sensing) (Tiaden et al 2007).

LqsR was found to promote phagocytosis, intracellular replication and cytotoxicity, and it plays a role in the transition from the stationary to the replicative growth phase (Tiaden et al 2007, Tiaden et al 2010). Additionally, it regulates the expression of virulence genes including 12 effector-encoding genes (Tiaden et al 2007). *L. pneumophila* strain lacking LqsA is highly defective for efficient uptake by phagocytes, while strains lacking LqsR shows various phenotypes regarding virulence (Tiaden et al 2010). The production of LqsR is controlled by RpoS (discussed below) and LetAS TCS (Tiaden et al 2010). Moreover, the production of LqsR was regulated at a posttranscriptional level by the sRNAs RsmY and RsmZ and by CsrA (Sahr et al 2009).

1.8.2 RpoS

RpoS (σ^S), is a sigma factor, a subunit of the RNA polymerase (Hengge-Aronis 2002, Trigui et al 2013). The σ^S factor has been extensively studied in multiple Gram-negative bacteria and it is important for the expression of virulence-related genes in several pathogens (Iriarte 1995, Merrell 2000, Hovel-Miner 2009). In *E. coli*, σ^S is considered the general stress sigma factor because it

senses and responds to different stress signals by regulating the transcription of multiple resistance and survival systems (Horwitz et al 1983).

In *L. pneumophila*, RpoS is involved in the regulation of virulence, stress response, and intracellular multiplication (Abu-Zant et al 2006, Hovel-Miner et al 2009). *L. pneumophila* strains lacking *rpoS* gene, which encodes σ^S, exhibited dramatic decrease in the intracellular multiplication within protozoan hosts (Hales et al 1999).

In an extensive study, Hovel-Miner et al. compared the global gene expression of an *rpoS* mutant to *L. pneumophila* wild-type during E and PE phases in rich medium. They showed that RpoS regulates a number of virulence factors including many Icm/Dot effectors and is required for intracellular multiplication (Hovel-Miner et al 2009). Further, during PE phase, RpoS upregulated the transcription of many virulence-related genes. Additionally, RpoS affects the expression of genes associated with multiple pathways required for *L. pneumophila* intracellular multiplication such as the transcription of the *cpxR* and *pmrA* (Hovel-Miner et al, 2009). In addition, the virulence regulator RpoS also controls the expression of the small RNA RsmY and RsmZ by regulating the expression of *letS* (Hovel-Miner et al 2009, Rasis and Segal 2009).

1.9 Small Regulatory RNA (sRNAs)

Small noncoding RNAs (sRNAs) is a class of regulatory elements that are involved in posttranscriptional gene regulation in both prokaryotes and eukaryotes (Waters and storz 2009). sRNAs are major players in the regulation of virulence factors and other processes in bacterial pathogens via interaction with mRNA targets and RNA-binding proteins (Waters and Storz 2009, Papenfort and Vogel, 2010). Earlier studies reported that sRNAs were associated with the control of replication of the ColE1 plasmid and transposition of Tn10 (Simons and Kleckner 1983).This early study presented sRNAs that were encoded on mobile genetic elements, however in 1984, MicF was the first reported bacterial sRNA encoded on the chromosome (Mizuno et al 1984).

MicF, was discovered in *E.coli* and was found to inhibit the translation of the major outer membrane porin OmpF throughout direct interaction with its target mRNA (Mizuno et al 1984). Since then, different approaches have been used to discover sRNA and reveal their significant role in regulation of gene expression, these approaches including: bioinformatic tools, RNA sequencing, and tiled microarrays with full genome coverage (Waters and Storz 2009).

sRNAs are short RNA molecules (80-500 nucleotides) that typically do not encode proteins. sRNAs are found in both intergenic regions and on the complementary strand of mRNAs (Weissenmayer et al 2011). It has been reported that most of sRNAs modulate gene expression at the post-transcriptional level (Weissenmayer et al 2011). Two different ways show how sRNA can be produced: sRNAs are transcribed from their own promoter and their transcription stops at a Rho-independent terminator, in other words they can originate from primary transcripts, or they can originate from the processing of larger transcripts (Faucher and Shuman 2011). In bacterial pathogens, including *L. pneumophila,* sRNAs regulate the expression of virulence genes and genes involved in the stress response important for survival in the host (Romby et al 2006).

1.9.1 Identification of sRNA in *L. pneumophila*

Putative sRNA molecules expressed by *L. pneumophila* have been identified by both bioinformatic approaches as well as by deep RNA-sequencing from growth in liquid bacterial culture and inside *A. castellanii* or by identification of *Legionella* homologs of sRNAs in other bacteria (Weissenmayer et al, 2011). Three main studies have identified sRNA expressed by *L. pneumophila*. First, Faucher el al (2010) used an *in silico* approach to identify sRNAs by searching for the presence of Rho-independent terminators in intergenic regions (Faucher et al 2010). 143 sRNAs were predicted in this study; however, microarray experiments showed that only few were actively transcribed. Six sRNAs were confirmed by Northern blot and rapid amplification of cDNA ends (RACE), including a 6S RNA homolog, which will be discussed below (Faucher et al

2010). The Second study was performed by Weissenmayer et al (2011). RNA sequencing was used to identify sRNAs expressed by the *L. pneumophila* Philadelphia-1 strain during E and PE phases in AYE broth and during infection of amoeba (Weissenmayer et al 2011). Sahr et al (2012) have done the third study using the same method as Weissenmayer group to identify sRNAs expressed by the Paris strain during the E and PE phases in AYE (Sahr et al 2012). Most of the identified sRNA are uncharacterized, only a few sRNAs have been studied in detail and those will be discussed below (Table 2).

Table 2. sRNAs that have been identified and characterized in *L.pneumophila* (The position of the sRNA is given relative to *L. pneumophila* Philadelphia-1 genome or Paris genome).

Name	5' end	3' end	Size (nt)	regulator	Target	Reference
Protein-binding sRNAs						
RsmX	3397553	3397653	101	LetA, RpoS	CsrA	(Sahr et al 2012)
RsmY	7168	7059	110	LetA, RpoS	CsrA	(Sahr et al 2012)
RsmZ	1892720	1892592	132	LetA, RpoS	CsrA	(Sahr et al 2012)
6S RNA	951819	951673	147		RNAP	(Faucher et al 2010)
Cis-encoded sRNAs						
LpPnc0584	2854265	2854119	146	LetA, RpoS	*phoA*	(Sahr et al 2012)
Trans-encoded sRNAs						
Lppnc405	1911295	1911105	196			(Sahr et al 2012)
Lpr0035	1355695	1355444	251		*lpg1229*	(Jayakumar et al 2012)

1.9.2 Base pairing sRNA

The most common type of regulatory sRNA are base-pairing sRNAs (Trigui et al 2013). They are short, highly structured RNA molecules called antisense sRNAs as they are complementary, to some degree, to their target mRNAs (Faucher and Shuman 2011, Brantl 2007). sRNAs can up-regulate or down-regulate the expression of a gene. (Waters and Storz 2009).They can bind at or

near the ribosomal binding site (RBS) to prevent recognition by the ribosome and subsequent translation (Faucher and Shuman 2011, Figure 1A). sRNA-binding to the mRNA can also induce its degradation (Waters and Storz 2009, Figure 1A). On the other hand, they can increase the stability of the target or modify the secondary structure of the mRNA target so that the ribosomal binding site (RBS) is exposed allowing translation initiation (Figure 1B).

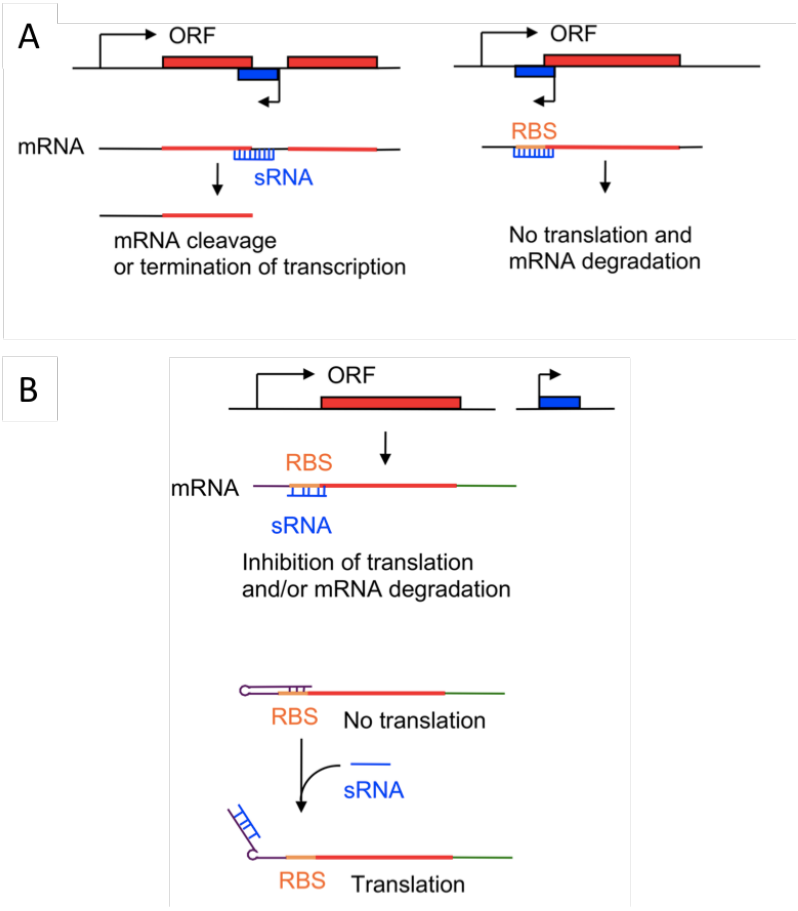

Figure 1. Mechanisms of sRNA-mediated regulation. A) Cis-encoded base pairing sRNA. B) Trans-encoded base pairing sRNAs. Source (Faucher and Shuman 2011)

1.9.2.1 Cis encoded sRNA

Cis-encoded sRNAs are located in the same DNA region on the complementary strand of their mRNA target and are, therefore, perfectly complementary to their mRNA targets (Figure 1A)

(Brantl 2007). Initially, the majority of cis-encoded sRNAs have been found in plasmids, phages and transposons (Brantl 2002). Because of technical difficulties to distinguish the source of expressed RNAs between the two DNA strands, the search for cis-encoded sRNAs using high throughput methods has been difficult compared to the search for *trans*-acting sRNAs (Cho and kim 2015). However, during the recent years, many cis-encoded sRNAs that are antisense to chromosomal genes have been discovered in bacteria using different approaches such as *in silico* or cDNA library sequencing (Cho and kim 2015). Cis-encoded sRNAs can control gene expression by binding to their mRNA targets, the binding can occur at the 5′end , 3′end, or in the middle of mRNAs (Cho and kim 2015, Trigui et al 2013). Depending on the location of their binding, regulatory mechanisms used by cis-encoded sRNAs varies between transcriptional attenuation, inhibition of translation or promotion of RNA degradation or cleavage (Brantl 2007). Thirty-three cis-encoded sRNAs were identified in *L. pneumophila* Philadelphia-1 that were at least partially complementary to genes encoding protein, some being known virulence factors (Weissenmayer et al 2011). Sahr et al identified around 450 cis-encoded sRNA in *L. pneumophila* strain Paris (Sahr et al 2012). Some cis encoded sRNA are antisense to genes encoding Icm/Dot effectors. For example, Lpr0050 is found antisense to the Icm/Dot effector SdeA and Lpr0003 and Lpr0004, are antisense to the gene encoding the Icm/ Dot effector LegA10 (*lpg2157*; Bardill et al, 2005, Weissenmayer et al 2011).

1.9.2.2 Trans encoded sRNA

Trans-encoded sRNAs are located in another chromosomal location than their targets and are only partially complementary to them (Figure 1B) (Brantl 2007, Trigui et al 2013). Trans-encoded sRNAs can regulate the expression of many different mRNAs. For instance, the sRNA RyhB in *E. coli* controls the expression of approximately 56 genes in response to changes in iron homeostasis (Massé et al 2005).

Because trans-encoded sRNAs are not fully complementary to their targets, they usually require the assistance of the RNA chaperone Hfq to enhance their interaction with their mRNA targets (Oliva et al 2017, Waters and Storz 2009). Hfq is a key regulator of expression of large variety of bacterial genes that impact many molecular processes in bacterial physiology, stress response, and virulence (Franze de Fernandez et al 1968, Updegrove et al 2016). In bacterial pathogens, deletion of the *hfq* gene often leads to a reduction in virulence, as was observed for *E. coli, Salmonella, Shigella, Yersinia,* and *Listeria* (Chao and Vogel 2010). In *Salmonella enterica*, Hfq is essential for optimal growth in epithelial cells as well as in macrophages (Sittka et al 2008). Additionally, in *Staphylococcus aureus,* deletion of *hfq* reduces virulence (Bohn et al 2007).

In *L. pneumophila,* the length of the lag phase period was affected in *hfq* mutant (McNealy et al 2005). Moreover, when a strain lacking *hfq* was grown in medium with low iron concentration, it exhibited reduction in the growth rate and reduced expression of the ferric uptake regulator (fur). In addition, the *hfq* mutant shows a small reduction in intracellular growth (McNealy et al, 2005). Therefore *L. pneumophila* base- pairing sRNAs acting through Hfq may regulate iron acquisition, virulence-related functions and possibly other systems. Around 38 trans-encoded sRNAs were identified in *L. pneumophila* Philadelphia (Weissenmayer et al 2011). Sahr et al identified around 91 trans-encoded sRNA in *L. pneumophila* strain Paris (Sahr et al 2012). More studies and more investigation are required to uncover their mechanism of action and their regulatory function (Cazalet et al 2004, Weissenmayer et al 2011).

1.9.3 Protein binding sRNA

Another class of sRNA influence gene expression by binding protein regulators of transcription (Waters and Storz 2009). RsmY, RsmZ and 6S RNA are the only sRNAs known to regulate gene expression through interaction with proteins (Trigui et al 2013, Sahr et al 2009).

1.9.3.1 CrsA and RsmYZ

CsrA is a RNA-binding protein first identified in *E. coli* where it was shown to control the expression of genes related to glycogen biosynthesis (Romeo et al 1993). An early study by Romeo (1998) reported that CsrA affects the stability or the translation of their mRNA targets by binding to a GGA motifs in the 5′UTR, effectively preventing translation (Romeo, 1998). In *E.coli*, CsrA expression is regulated by the sRNA CrsB and CsrC. These sRNAs contain many GGA motifs allowing them to bind multiple CsrA proteins resulting in sequestration of CsrA. This, in turn, prevent CsrA from binding to its target mRNAs (Romeo et al 1993). In *E.coli*, transcription of CsrB and CsrC is regulated by the BarA/UvrY TCS (Trigui et al 2013, Sahr et al 2009).

A similar regulatory system is found in *L. pneumophila*, where CsrA is responsible for the repression of transmissive traits during E phase, including pigmentation, motility, and cell shortening (Fettes et al 2001, Molofsky and Swanson 2003). Moreover, CsrA is required for intracellular growth in both mammalian macrophages and *A. castellanii* (Molofsky and Swanson 2003). In *L. pneumophila*, CrsA is regulated by the two sRNA RsmY/Z which are CsrB homologs (Kulkarni et al 2006). In *L. pneumophila*, expression of RsmY/Z is regulated by the LetA/S TCS (Faucher and Shuman 2011, Molofsky and Swanson 2003).

When infecting a host cell, it is essential that *L. pneumophila* turn off survival and transmission genes, whereas genes important for intracellular replication are turned on (Trigui et al 2013). This is regulated by LetAS-RsmYZ-CsrA regulatory cascade (Segal 2013, Trigui et al 2013). During E phase CsrA represses the expression of transmissive phase genes (Shar et al 2010). Direct regulation by CsrA is post-transcriptional and occurs by CsrA binding to the mRNA of target genes (Babitzke and Romeo, 2007).

During PE phase LetA/S activates the expression of RsmY/Z (Sahr et al 2009) which will bind to CsrA (Trigui et al 2013) allowing expression of post-exponential traits (pigmentation, cytotoxicity,

and motility). RsmY/Z are essential for the expression of some virulence determinants, and for the switch between the replicative and transmissive phases of *L. pneumophila* (Trigui et al 2013). Deletion of either *rsmY* or *rsmZ* had minimum effect on virulence, but deletion of both strongly reduced replication in macrophages as well as in *A. castellanii* (Sahr et al., 2009). Some Icm/Dot effector genes (*ralF* , *sidC* , *sdeA* , *sdeC* , *sidF* , *sdhB* , *legC7* , *legC2*, and *vipA*) are regulated by LetA/S in an RsmY/Z-dependant manner (Rasis and Segal 2009, Sahr et al 2009).

1.9.3.2 6S RNA

The 6S RNA was first identified and sequenced in *E. coli* 40 years ago (Hindley 1967). Wassarman and Storz (2000) showed that 6S RNA binds to σ^{70}-containing RNA polymerase (RNAP) and acts to reduce its activity in stationary phase. (Wassarman 2007, Cavanagh et al 2008). In 2010, Faucher et al (2010) identified 6S RNA in *L. pneumophila* based on its expression pattern and predicted secondary structure which is similar to the *E. coli* 6S RNA, and by its capacity to bind to *L. pneumophila* RNAP (Faucher et al 2010). A second 6S RNA, named 6S2 RNA was identified in the Philadelphia-1 strain along with an antisense transcript (Weissenmayer et al. 2011); however, only the sense transcript was detected in the Paris strain (Sahr et al 2012). The deletion of 6S RNA significantly reduced *L. pneumophila* intracellular multiplication in macrophages as well as in *A. castellanii* (Faucher et al 2010). Strain lacking 6S was also unable to compete against the wild-type strain during infection of macrophages and *A. castellanii*. Therefore, 6S RNA is essential for optimal intracellular growth (Faucher et al 2010, Trigui et al 2013).

1.10 *Legionella* genome

Genome sequencing is required in order to understand the genetic characteristics of *L. pneumophila* (Cazalet et al 2004). Several strains of *L. pneumophila* have been sequenced, for example; Philadelphia-1 (Chien et al 2004) derived from the original Philadelphia outbreak (McDade et al 1977), Paris (Cazalet et al 2004), Lens (Cazalet et al., 2004), Alcoy (D'Auria et al

2010), Corby (Glockner et al 2008), Lorraine (Gomez-Valero et al 2011), HL 0604 (Gomez-Valero et al 2011) and 570-CO-H (Amaro et al 2012). The analysis of *L. pneumophila* genomes provide a reference for molecular epidemiology as well as allow researchers to better understand the fundamental mechanisms of pathogenesis of *L. pneumophila* (Newton et al 2010). Moreover, genome analysis as well as expression profiling of *L. pneumophila* and its host are required to identify various regulatory pathways that mediate the adaptation of *L. pneumophila* to the different intracellular environment (Jules and Buchrieser 2007)

The general features of the genomes including size (3.3 to 3.5 Mb), GC content (38%), number of predicted genes (3,001 to 3,259), percentage of coding regions (88 to 90.2%), and average length of the coding regions are conserved amongst isolates (Cazalet et al 2004, Chien et al 2004, Glockner et al 2008).

Bacteriophages, plasmids and other mobile genetic elements (MGE) are essential players in genome plasticity (Dagan et al 2008). Analysis and comparison of the genome sequences of the Paris and Lens strains of clinical *L. pneumophila* revealed that this bacterium is highly versatile organism that shows extensive genome plasticity and diversity (Cazalet et al 2004). In addition, some strains contain chromosomal regions, which can be excised and maintained as plasmids-like elements (Cazalet et al 2004, Chien et al 2004, Doleans et al 2004). These elements contain a type IVA secretion system: the Lvh T4SS in strains Paris, Lens and Philadelphia-1 and two new type IVA secretion systems (trb/tra) in strain Corby (Glockner et al 2008). The Lvh type IVA secretion system was reported to be involved in virulence (Jules and Buchrieser 2007) thus diversity in this system as well as differences in copy number due to excision might be of importance for survival and spread of *L. pneumophila* in the environment (Doleans et al 2006). Overall, the excision and

integration of plasmids or genes might represent one mechanism that *L. pneumophila* exploits to adapt to different environments (Cazalet et al 2004).

2. Hypothesis and objectives

In recent years small noncoding RNAs (sRNAs) have become a major player of posttranscriptional gene regulation in both prokaryotes and eukaryotes (Oliva et al 2017). sRNAs have many regulatory roles including stress response, virulence, and competence (Trigui et al 2013). About thirty-three cis encoded sRNAs and around 38 trans encoded sRNA have been identified in *L. pneumophila* Philadelphia-1 but the majority have not been previously characterized (Weissenmayer et al 2011). The main objective of my research project is to characterize two of these sRNAs, the cis-encoded sRNA Lpr0050 and the trans-encoded sRNA Lpr0024 (Table 3).

In general, we hypothesized that these two sRNAs play role in the regulation of virulence trait in *L. pneumophila.*

Table3. Lpr0050 and Lpr0024 (The position of the sRNA is given relative to *L. pneumophila* Philadelphia-1

Name	5' end	3' end	Size (nt)	Target	Reference
Lpr0050	2418574	2418506	68	lpg2157	(Weissenmayer et al 2011)
Lpr0024	816705	816590	115	Unknown	(Weissenmayer et al 2011)

Lpr0050 is encoded complimentary to Icm/Dot substrate SdeA. Therefore, our hypothesis is that Lpr0050 regulates the expression of *sdeA* and its homologs. The objectives are:

- Confirm expression of lpr0050 by northern blotting.
- Study the role of lpr0050 by cloning *sdeA* fragment complementary to lpr0050 then measuring the expression of lpr0050 by northern.
- Test the effect of lpr0050 on the expression of *sdeA* and its homologs by qPCR.

Lpr0024 is a trans-encoded sRNA and its target is therefore unclear. It is found in multiple copies in the genome. Hypothesis: the genomic region of Lpr0024 can excise from the chromosome and exist in an episomal form

Objectives

- Confirm expression of lpr0024 by northern blotting
- Test if there is recombination leading to the excision of that region, by PCR and Sanger sequencing

3. Materials and methods

3.1 Media, antibiotics and growth conditions:

L. pneumophila was routinely grown on CYE (charcoal yeast extract) agar supplemented with 0.1 mg /ml α-ketoglutarate, 0.25 mg /ml L-cysteine and 0.4 mg /ml ferric pyrophosphate (Feeley et al, 1979). Plates were incubated at 37 °C for three days. For liquid culture, AYE (ACES buffer yeast extract) broth was used (Sadosky et al, 1993). *L. pneumophila* was inoculated in AYE at an optical density of 600 nm (OD_{600}) of 0.1 overnight at 37°C with shaking. Exponential phase (E) bacteria were harvested at an OD_{600} of 0.7–0.8 and post-exponential phase (PE) bacteria were harvested at an OD_{600} between 3.5 and 4.0. Antibiotics and supplements were used at the following concentrations: kanamycin, 25 μg/mL; chloramphenicol, 5 μg/mL; IPTG, 1 mM. *E. coli* strains were grown on LB plates and in LB broth (Sezonovo et al 2007). If needed, either chloramphenicol was added to a final concentration of 25 μg/mL, or ampicillin at 100 μg/mL.

3.2 Bacterial strains and plasmids

All strains used in this study are listed in Table4. Strains were stored at −80 °C in AYE with 10% glycerol. All *L. pneumophila* strains are derivatives of JR32, a streptomycin resistant, salt-sensitive and restriction negative mutant of *L. pneumophila* strain Philadelphia-1 (Sadosky et al 1993). KS79 is a Δ*comR* mutant of the JR32 strain making it actively competent, which in turn makes allelic exchange mutations through natural transformation possible (de Felipe et al 2008). Therefore, KS79 will be referred to as the wild-type (WT) strain in this study.

3.3 Cloning of the complementary sequence to *lpr0050*

A 300 bp fragment of *sdeA* complementary to the sRNA *lpr0050*, was cloned into pMMB207c plasmid under *Ptac* promoter (Morales et al, 1991). The fragment was amplified from the KS79 wild-type strain by PCR using One Taq *polymerase* (NEB) according to the the manufacturer's protocol. The annealing temperature used was 54°C. Primers trap-F and trap-R (Table 5) were

used which contain EcoRI and XbaI restrictions sites respectively (underlined in Table 52). PCR amplicon was visualized on 0.7% agarose gels. The amplicon and the pMMB207c plasmid were then digested using EcoRI and XbaI (NEB) according to the manufacturer's protocol. The digested amplicon and the digested plasmid were then ligated using T4 DNA ligase (NEB). The recombinant plasmids were then transformed into *E. coli* DH5α competent cells which were made according to CCMB protocol. Colonies that grew on antibiotic plates (LB plates with 25ug/ ml chloramphenicol) were patched and tested by PCR to confirm the insertion of the insert into pMMB207c plasmid. The resulting plasmid was named pSF118 (Table 4). The plasmid was extracted using a plasmid extraction kit (Qiagen) according to the manufacturer's protocol and introduced into the wild type strain KS79 by electroporation (Chen et al 2006) to produce strain SPF482.

3.4 RNA Extraction

Strains were grown to E-phase and PE-phase as explained above. To extract RNA, 10 mL of exponential phase culture or 5 mL of post exponential phase culture was centrifuged for 10 minutes at 4500 rpm. Then the supernatant was discarded and the pellet was suspended in 1mL of TRIzol reagent (Invitrogen). RNA was extracted according to the manufacturer's protocol. For quantitative PCR (qPCR) analysis, the extracted RNA was treated with TURBO DNase (Invitrogen) and incubated for 30 minutes at 37 °C. The DNase was inactivated by adding DNase inactivation buffer (Invitrogen). The purity and quantity of the RNA were measured using a NanoDrop (Thermo Scientific).

3.5 quantitative PCR (qPCR)

RNA was extracted from WT and from SPF482 ON and OFF as explained above. cDNA was prepared using 1μg of total RNA (Protoscript II, NEB) following the manufacturer's instructions. qPCR was performed according to the manufacturer's protocol using the iTaq Universal SYBR

Green Supermix (Bio-Rad) and primers in (Table 5). For each strain, three replicates of cDNA were used and a no reverse-transcriptase reaction was carried out as a negative control. For each tested gene, two sets of primers were designed; one set was designed at the 5' end of the gene (before the binding site) and the other after at the 3' end after the (binding site). 16S rRNA was used as a house keeping gene. A relative quantification strategy was used to perform analysis of the qPCR data. The calculated CT was normalized to the value of the house keeping gene control 16S amplified from the corresponding samples and the fold-change of the gene expression was calculated as previously described (Livak and Schmittgen, 2001).

3.6 Northern Blotting

Strains were grown to exponential and post exponential phases and RNA was extracted as described above. Samples were prepared in PCR tubes by combining 1 or 10 µg of RNA with 5 µl of Novex TBE Urea sample buffer (Invitogen). For the ladder, 2 µl RNA century plus marker (Invitogen) was mixed with 2 µl of Novex TBE Urea sample buffer. Samples were heated at 70 °C for three minutes to denature the samples. RNA was loaded and migrated on 6% polyacrylamide urea gel at 180 mV in TBE buffer. The RNA was transferred onto a positively charged nylon membrane (Thermo Scientific) using a semi-dry gel blotting system (BioRad) for 20-30 minutes at 200 mA. The membrane was placed into UV Stratilinker 1800 in order to fix RNA to the membrane. The membranes were pre-hybridized in ULTRAhyb-Oligo Hybridization Buffer (Ambion) for 1 hour at 37 °C before hybridization with the designed oligonucleotides (60-mers) coupled to biotin at the 5' end Table 4. Hybridization was done overnight in a rotating chamber at 37 °C. Membranes were washed twice with 2X SSC (0.15 M NaCl and 0.015 M sodium citrate) and 0.5% SDS for 30 minutes at 37 °C in a rotating chamber. Detection was performed using the Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific) as per the manufacturer's instructions.

3.7 DNA Sequencing

Analysis of the genomic region of lpr0024 was performed from the *L. pneumophila* Philadelphia-1 strain by PCR using One Taq polymerase (NEB). Reaction set up and thermocycling conditions were as per the manufacturer's protocol, annealing temperature used is 56 °C. The primers used are described in (Table 5) and shown in (Figure 8). Primers iraB-F and yhbQ-R were used to amplify the region between *iraB* and *yhbQ* and are shown in blue arrows Primers episome-R and episome-F were used to test the episome formation and are shown in green arrows. PCR amplicons were visualized on 0.7% agarose gels. PCR products were ligated into pGEM-T Easy Vector (Promega) following the manufacturer's instructions using an overnight incubation at 4 °C. The ligation products were transformed into *E. coli* DH5α chemically competent cells as per CCMB protocol. Transformants were selected on LB agar plates containing 100 µg/mL Ampicillin, 0.5 mM IPTG and 80 µg/mL X-Gal. The recombinant plasmids were extracted using a plasmid extraction kit (Qiagen) according to the manufacturer's protocol. The purity and quantity of the DNA were measured by NanoDrop. In order to confirm insertion of the correct fragment, restriction digestion reaction was done using EcoRI (NEB), reactions were incubated at 37°C for 1 hour. Digestion products were visualized on 0.7% agarose gel. Sequencing of the amplicons was performed at the Plate-forme d'Analyses Génomiques de l'Université Laval, Quebec City, Quebec, Canada.

Table 4. Strains and plasmids

Strain name	Relevant genotype	Reference
L. pneumophila		
JR32	Philadelphia-1; Sm^r; r^- m^+	(Sadosky et al, 1993)
KS79 (WT)	JR32 Δ*comR*	de Felipe et al. (2008
Δ*letS* (GAH338)	KS79 *letS*:: Kn^r	**(**Hovel-Miner et al 2009)
Δ*cpxR* (SPF159)	KS79 cpxR::Kn^r	Lab collection
SPF482	Ks79+pSF118, Cm^r	This study
E. coli		
DH5α	F^- endAl hsdR17 (r^- m^+) supE44 thi-1 λ^- recAl relAl Δ(argF-lacZYA)U169 ϕ80dlacZΔM15 deoR gyrA96 Nal^r	NEB
Plasmids		
pMMB207c	RSF1010 derivative, IncQ, $lacI^q$ Cm^r P*tac* *oriT* Δ*mobA*	Morales et al, 1991
pSF118	pMMB207c + SdeA (only region complementary to lpr0050 was cloned , Cm^r	This study
pGEMT-easy	PCR cloning vector, Amp^r	Promega

Sm^r, streptomycin resistance; Cm^r, chloramphenical resistance; Gm^r, gentamicin resistance; Km^r, kanamycin resistance.

Table 5. Primers and oligonucleotides probes

Name	Sequence
Trap F	TCATCA GAATTC GAGCACGATGCGAAAACGATG
Trap R	TCATCA TCTAGA TGTCGGTGTCATTGAGGAGTG
sdeA 5end F	GCACGCATGGGGTATGGAGAT
sdeA 5end R	ACACGGGCATAAAACCTTGCC
sdeA 3end F	GAACTGGATCGGCTGGACAGA
sdeA 3end R	TGTCAAGTCGTTCGTCTGTGG
sdeC 5' F	GGAGTACCGACTATTGACACAGAG
sdeC 5' R	GATGTTTTACCCAGTGGGTCAAAG
sdeC 3' F	GCTGAAGCAAGAAGAGTGAGACTTC
sdeC 3' R	CGCTCTAAGAGTTTGCTCCTTC
sdeB 5' F	GGCATCCCGAGATTAAAAGCG
sdeB 5' R	GAGTCTCTCTTAGGTCCTACCC
sdeB 3' F	CGCCTGTAGAAATTCTCTCCTTG
sdeB 3' R	GATTATCAAGTGAGTGAAGCCTGG
sidE 5' F	GGATGCATGCTATTTTTGAGCGC
sidE 5' R	GGCATAAAGCCTTGCCTGTCAAG
sidE 3' F	GAAGCTCATCAAGCTCAAATGGC
sidE 3' R	CAAGAGTTACCCGCTCATCTGTAG
sidJ 5' F	CGGTGGAAACAACCGATGTTTC
sidJ 5' R	GATTGTCTCGGGTTTGACAGAACC
sidJ 3' F	CAGGACTTGGGTGGATATAACAGG
sidJ 3' R	CTGCTTCAAAGCACGAATTGGC
iraB_F	GCTTTGGGGCTTGGTATTCTGATG
yhbQ_R	GTCTGATAACGCTTGCCAGTCTCC
episome-R	GCCCTGTATAATAACTCTTGTCAC
episome-F	GGCAAGCGTTATCAGACTATTC
lpr0024 probe	CCCATCGGGCCATTCTTGCTTCAAGATGGTGCTTCGCACTTCCTCAGCAT GAACGGT

lpr0050 probe	CGATGAAATTGCTGCTGCCCGCATCATTCAGCAAATCTTGGCCAATCCTG ATTGCAT
Trap probe	ATGCAATCAGGATTGGCCAAGATTTGCTGAATGATGCGGGCAGCAGCAA TTTCATCG

XbaI and EcoRI restriction sites are underlined

4. Results

4.1 *lpr0050* is expressed in the E and PE phases

The cis-encoded *lpr0050* was detected by Northern Blot (Figure 2). Results revealed that *lpr0050* is expressed in both E and PE phases in the KS79 wild-type strain. Furthermore, it is also expressed in Δ*letS* and Δ*cpxR* at the wild-type level, indicating that *lpr0050* is not regulated by LetS nor CpxR.

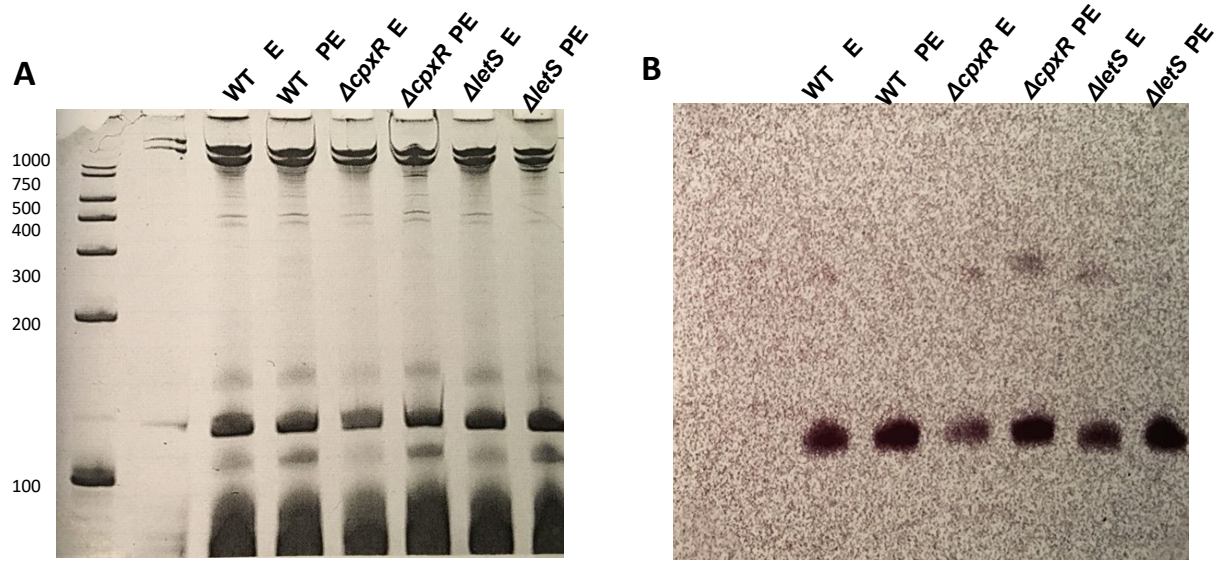

Figure 2. Expression of *lpr0050* by northern blot in exponential (E) and post exponential (PE) phases in WT, *ΔcpxR* and *ΔletS* strains. A) RNA was loaded and migrated on 6% polyacrylamide urea gel. RNA century plus marker was used. The gels were stained with ethidium bromide (EtBr) and viewed under UV light. B) Detection of Lpr0050 by using oligonucleotides (60-mers) coupled to biotin at the 5'end.

2.4 A fragment of *sdeA* complementary to lpr0050 is overexpressed in the E phase

lpr0050 is encoded on the complementary strand of the Icm/Dot effector SdeA (*lpg2157*; Bardill et al 2005). Therefore, *lpr0050* cannot be deleted without affecting its target. In order to study the role of the sRNA, a fragment of *sdeA* complementary to the sRNA, called the trap, was cloned. This method was used by Duhring et al to study the role of the cis encoded sRNA IsrR which regulates expression of the photosynthesis gene *isiA* in the cyanobacterium *Synechocystis* sp. PCC

6803 (Duhring et al 2006). The rationale is that overexpression of the trap could block lpr0050 function by sequestering it and therefore preventing it from interacting with the target, *sdeA*. In order to test this hypothesis, the trap, which is 300bp, was cloned in pMMB207c under the *Ptac* promoter and transformed into *L. pneumophila* to produce SPF482 (Table 4). Northern blotting was performed to test the expression of the trap using 5′ biotinylated Trap probe (Table 5). The expression was tested in exponential (E) phase and in post exponential (PE) phase in two strains: KS79 wild-type and SPF482 grown with 1mM IPTG to induce the expression of the trap (ON) or without IPTG (OFF) (Figure 3 A, B). The results showed that the 300 bp trap is expressed only in the E phase in the (ON) condition. It is not expressed in the PE phase maybe because it is degraded but the absolute reason is unknown.

In order to test if overexpression of the trap affects the expression level of *lpr0050*, northern blot was performed to test the expression of the sRNA after the trap was overexpressed (Figure 3 C , D). The results demonstrate that *lpr0050* is expressed normally, which suggests that overexpression of the trap does not affect the stability of the sRNA. However, it could still affect or block its function. Therefore, we moved forward to test this by measuring the expression of *sdeA* and its homologos.

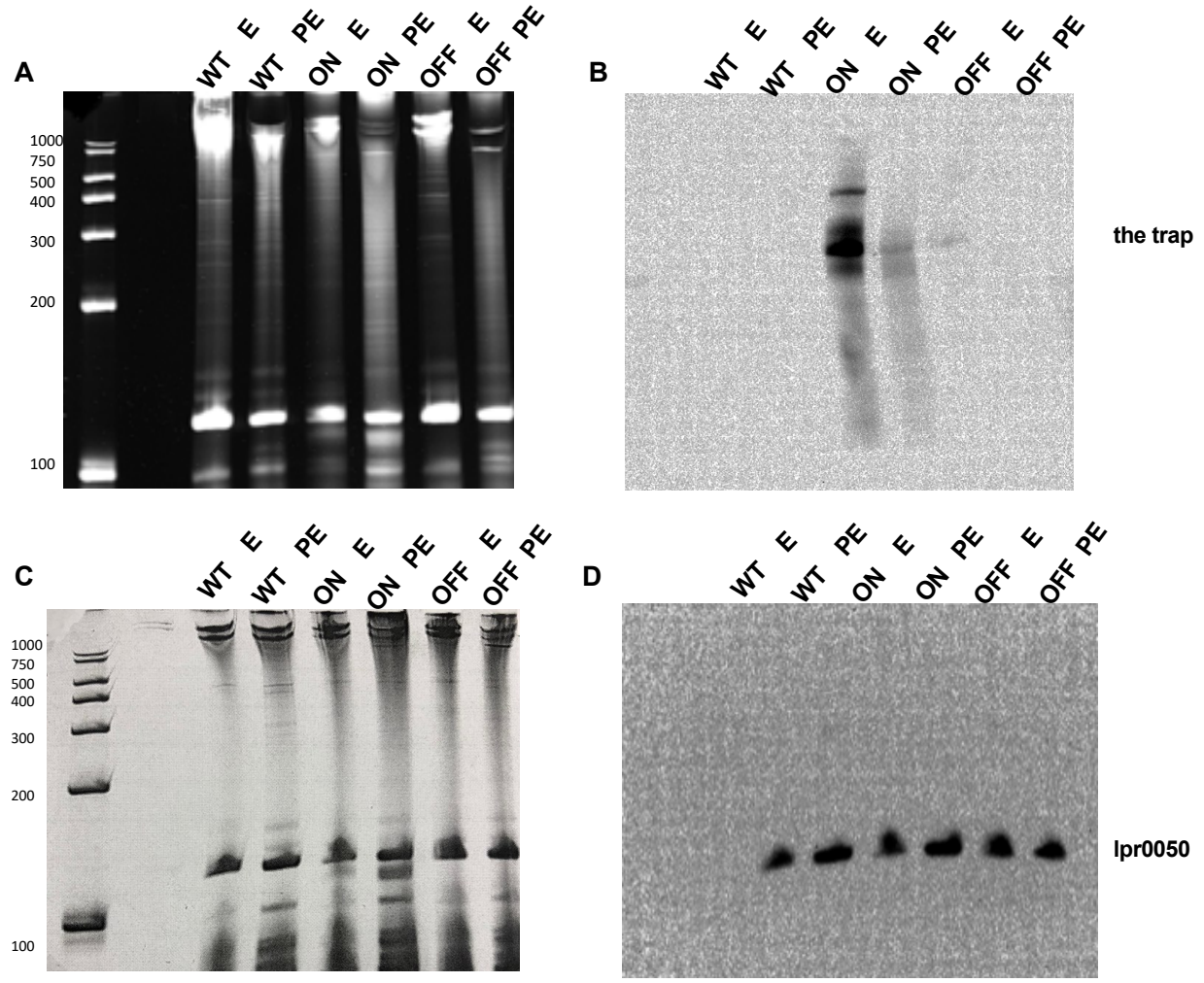

Figure 3. Expression of the trap by NB (A, B) and *lpr0050* (C, D) was tested by northern blot. in exponential (E) and post exponential (PE) phases in KS79 wild-type strain, and in SPF482 with (ON) and without (OFF) IPTG. A&C) Gel electrophoresis. RNA was loaded and migrated on 6% polyacrylamide urea gel. RNA century plus marker was used. The gels were stained with ethidium bromide (EtBr) and viewed under UV light. B&D) Detection of the trap (C) and lpr50 (D) using oligonucleotides (60-mers) coupled to biotin at the 5' end.

3.4 Lpr0050 regulates the expression of *sdeA* and its homologs

Overexpression of the trap had no effect on *lpr0050* expression (Figure 3 C, D), one explanation is that the trap could block the function of the sRNA by titration/sequestration which in turn could affect the expression of *sdeA*. To test this, the expression of *sdeA* was measured by qPCR (Figure 4 B). Two sets of primers were designed (Table 5) one set of primers was designed at the 5' end

of *sdeA*, whereas the other set was designed at the 3' end. The rational is to measure the expression of the transcript before and after the lpr0050 binding site. The results show that for the beginning of the transcript (5' end) there is no significant difference in the expression of *sdeA* whether the trap is turned on or off. In contrast, expression of the end of the transcript (3' end) is reduced when the trap is expressed. This result suggests that the trap can block the function of *lpr0050* and therefore affect the expression of the mRNA target (*sdeA*). Which suggests that *lpr0050* regulates the expression of *sdeA*.

Since SdeA shares extensive homology with the other SidE homologs, it was hypothesized that lpr0050 could regulate their expression as well. The putative binding site in the other SidE homologs was predicted using RNA predator (Figure 5) (Florain et al 2011).To test if the other SidE members are affected by expression of the trap, their expression was also measured using the same strategy as for *sdeA* (Figure 4 B). Overexpression of the trap induces the expression of the whole transcript of the other *sidE* homologs. The expression of *sidJ* was also measured (Figure 4 B). SidJ is located in the same locus but not part of the SidE family. The results reveal that expression of *sidJ* is decreased upon the expression of the trap, and contrary to *sdeA*, the whole transcript is affected.

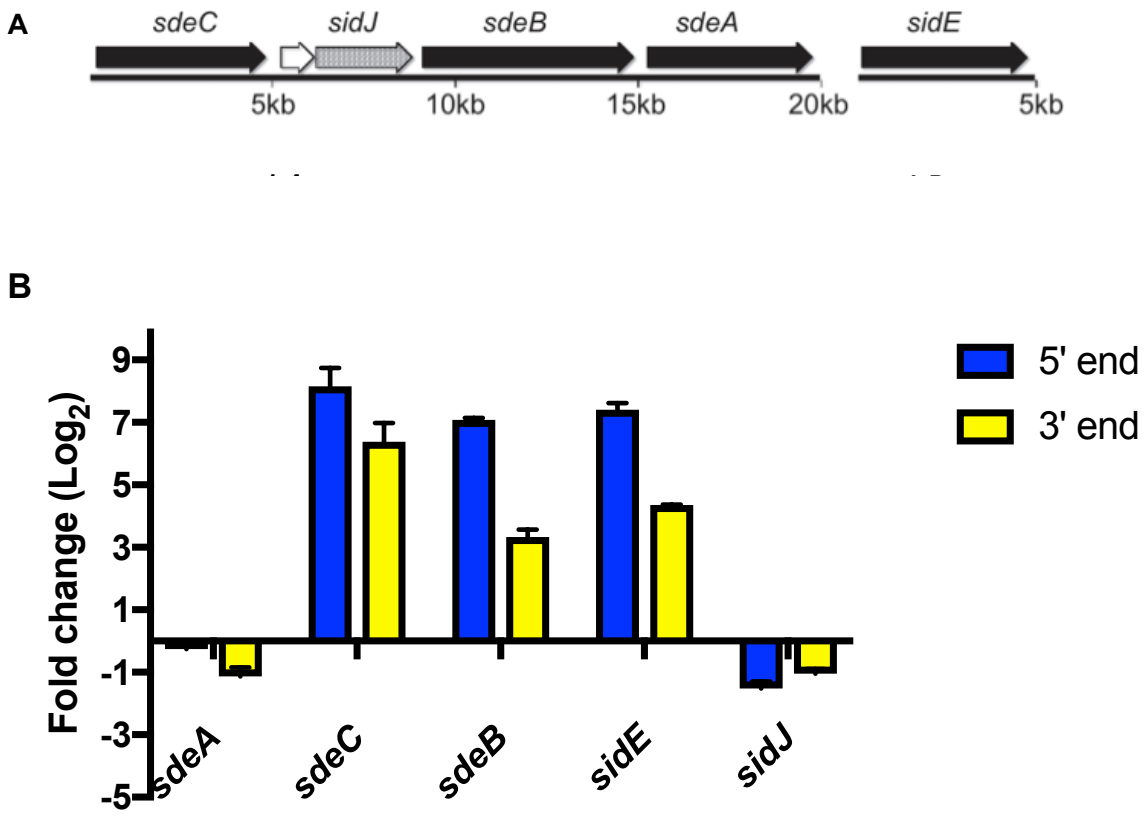

Figure 4. qPCR was used to measure the expression of sidE family genes in E phase. A) Depiction of the genetic organization of SidE family effector and SidJ (adapted from Jeong et al 2015). *sdeA, sdeB, sdeC* are located in *sdeC-sdeA* locus whereas *sidE* is located on different location. *sidJ* is located on the same locus but it is not part of the SidE family. B) qPCR was used to measure the expression of *sdeA*, the mRNA target of *lpr0050* as well as the expression of the *sdeA* homologs *sdeB, sdeC, side sidJ.* The results compare SPF482 grown with IPTG (ON) to SPF482 grown without IPTG (OFF). The OFF fold changes are not statistically significant.

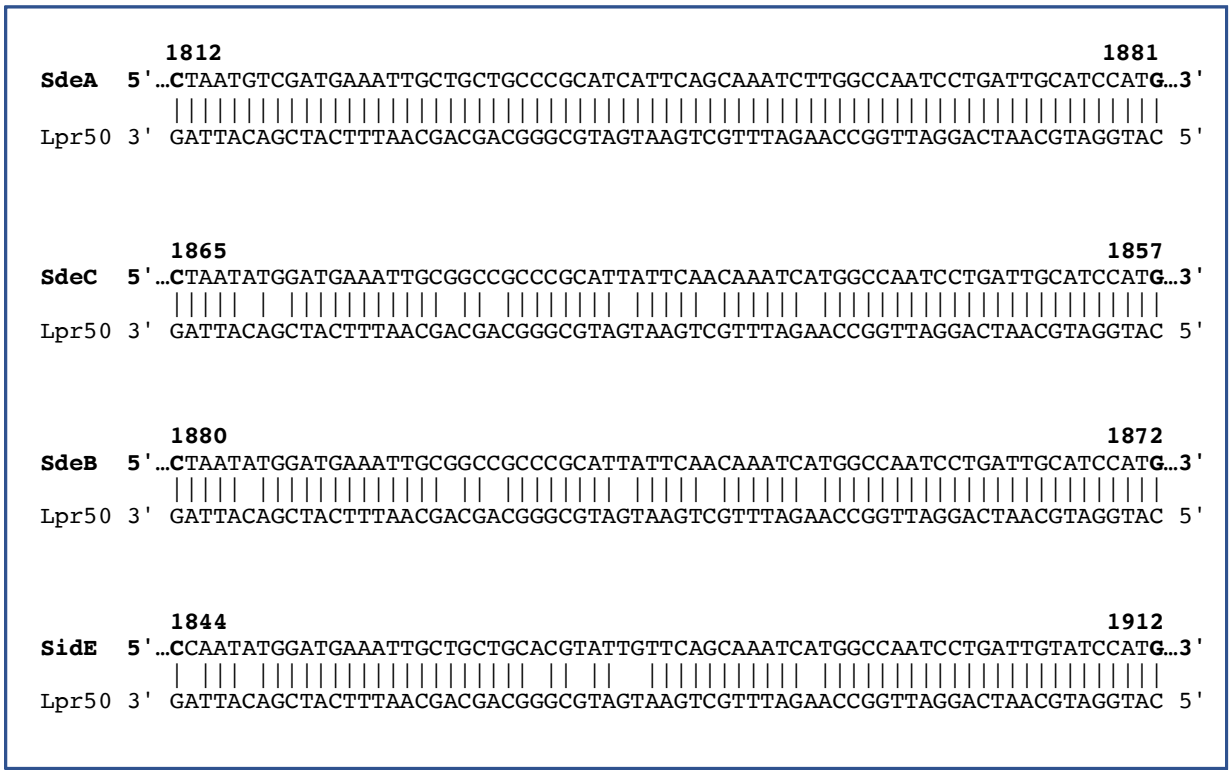

Figure 5. Corresponding sequence and complementarity between Lpr0050 and SdeA and its homologs. The putative binding site was predicted using RNA predator.

4.4 *lpr0024* is regulated by LetS in the PE phase

The expression of *lpr0024* was detected by Northern blot (Figure 6 A , B). The results showed that *lpr0024* is expressed in both E and PE phases in the KS79 wild-type strain. Deletion of *cpxR* resulted in wild-type expression in E phase but reduced expression in PE phase. Similarly Lpr0024 is only expressed in the E phase in Δ*letS*. These results indicate that *lpr0024* is regulated by CpxR and LetS in the PE phase.

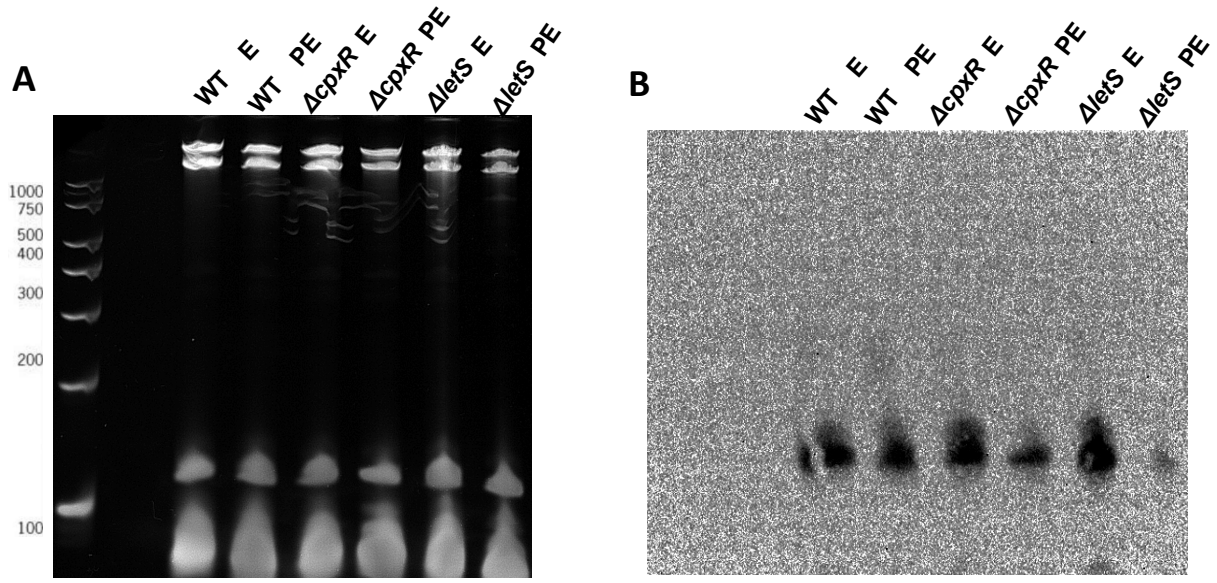

Figure 6. Expression of *lpr0024* by Northern blot in exponential (E) and post exponential (PE) phases in WT, *ΔcpxR* and *ΔletS*. A) RNA was loaded and migrated on 6% polyacrylamide urea gel. RNA century plus marker was used .The gels were stained with ethidium bromide (EtBr) and viewed under UV light. B) Detection of *lpr0050* by using oligonucleotides (60-mers) coupled to biotin at the 5'end.

4.5 Lpr0024 is a multicopy sRNA

Figure 7 shows the analysis of the genomic region of the trans encoded sRNA lpr0024. The goal was to design primers to delete *lpr0024.* After extensive inspection of this region it was clear that the lpr0024 sequence was repeated multiple times. Detailed analysis revealed that *lpr0024* is encoded between *iraB* and *lpg0745* genes and it is present in several copies in the *L. pneumophila* strain Philadelphia-1 genome.

Additionally, a previously unannotated ORF at the 3' side of the sRNA was identified. Using BLAST analysis, the ORF was identified as a putative endonuclease that belongs to GIY-YIG nucleases superfamily(Kumar et al 2016). We found that that ORF is homologous to *yhbQ* which is found in *E. coli* and other bacteria, therefore we named the ORF *yhbQ*. The analysis revealed that there are four full copies of *lpr0024* and two partial copies are encoded on both sides of *yhbQ*.

The precise sequences of the full and partial copies was double-checked to ensure they correspond to the genomic sequences, which have slight differences, and not copies produced by repriming of PCR amplicons on themselves.

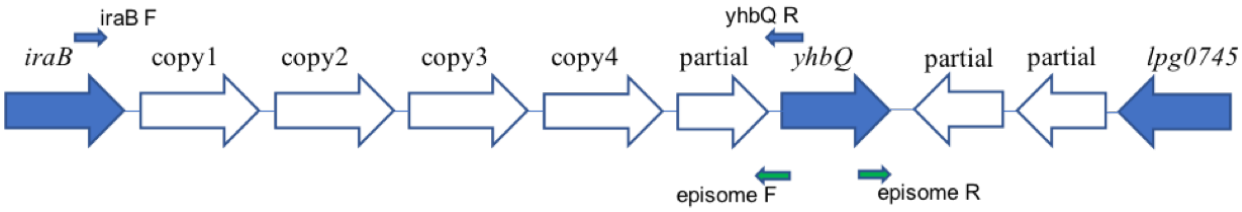

Figure 7. Analysis of the genomic region of *lpr0024*. *lpr0024* is encoded between *iraB* and *lpg0745* and it is present in several copies. Primers used in this study are: iraB-F & yhbQ-R (blue arrows **)** for PCR analysis of the number of copies (Figure 8) , and episome-F & episome-R (green arrows) for PCR analysis of the formation of an episome (Figure 9).

4.6 Recombination between the repeated sequences leads to variation of lpr0024 copy numbers

The repeated sequences suggest possible recombination between them which will results in a variation in the copy numbers of the sRNA at the population level. To test this hypothesis, PCR analysis of the genomic region between *iraB* and *yhbQ* was performed (Figure 8). The expected size of the amplicon with the four full copies is 1172 bp: 632 bp for the four full copies (158 bp each) and 540 bp for the rest of the region including *yhbQ*. As the results show, the PCR resulted in multiple amplicons of different size suggesting that the number of copies of the sRNA actually varies at the population level. For instance, an amplicon of the size of 1200 bp contains the four full copies. Some bands were stronger, such as the bands at 700 bp (one copy) and the one at 800 bp (two copies), indicating that some number of copies are favored at the population level. To confirm these results, the amplicons were cloned in pGEM-T easy vector and sequenced by Sanger sequencing. The sequencing results are represented in (Figure 8 and Table 6).

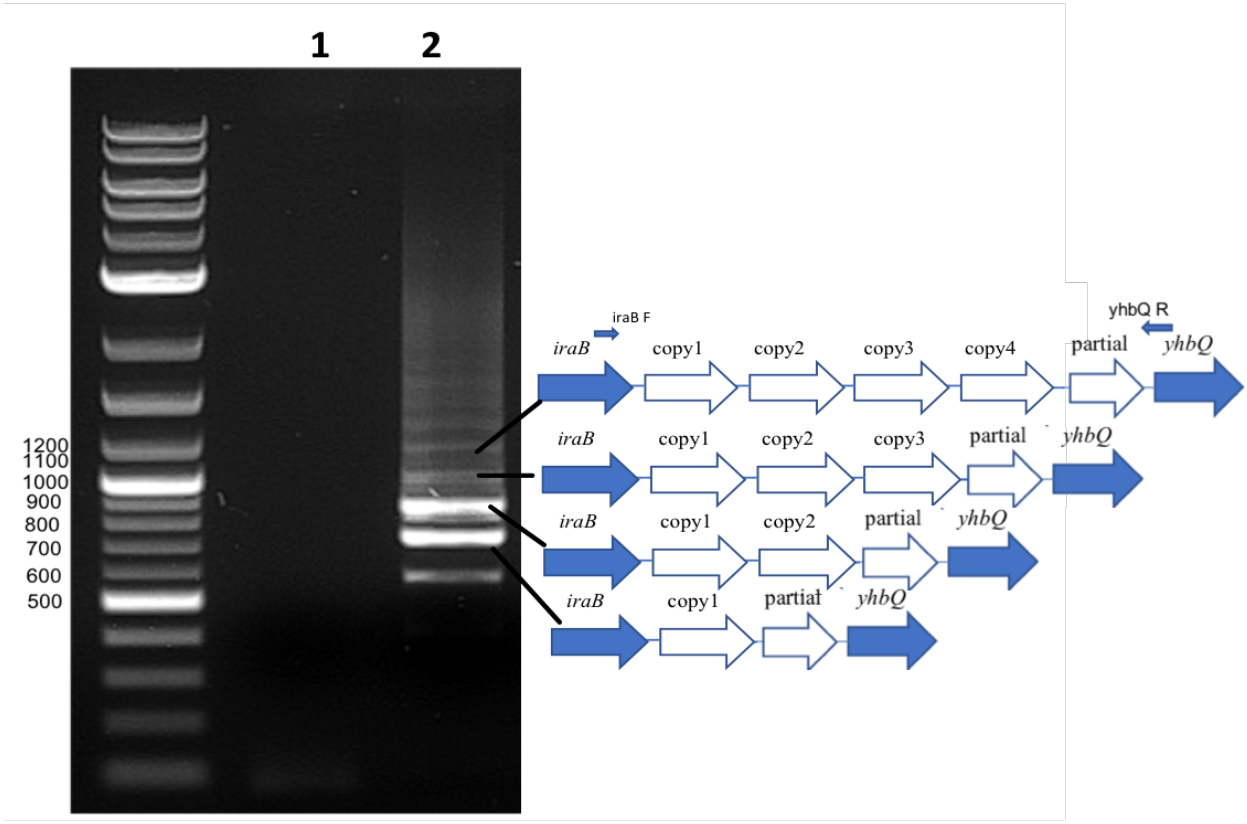

Figure 8 . PCR analysis and sequencing of Lpr0024 genomic region. The size of the genomic region between *iraB* and the *yhbQ* was investigated by PCR. Samples are: 1, negative control (water); 2, DNA from the wild type. Primers iraB-F and yhbQ-R were used. The genomic organization of some of the bands is depicted on the right side. The expected size of the amplicon containing 4 copies is 1172 bp. Amplicons were cloned with pGEMT-easy and were sequenced by Sanger sequencing to confirm the number of copies of the sRNA (Table 5).

Table 6. Sequencing results of the region between *iraB* and *yhbQ*

Colony #	Size on Gel	Size based on sequencing	# copies
2	~800 bp	856 bp	2
3	~1100 bp	1172 bp	4
4	~700 bp	698 bp	1
5	~1000 bp	1014 bp	3
6	~1200 bp	1330 bp	5
8	~1100 bp	1172 bp	4
9	~1100 bp	1172 bp	4

4.7 Lpr0024 can exist in an episomal form

The recombination event between the repeated sequences with any of the two inverted partial copy downstream of *yhbQ* could lead to the excision of this region from the chromosome. To test this hypothesis, PCR analysis was performed using primers episome-F and episome-R (shown in green arrows in Figure 7). These primers amplify the region between the partial copies encoded on both sides of *yhbQ* that presumably exists in the episome to investigate the presence of the episomal form. If an episome is formed, we would get an amplicon which represents recombination between the repeated sequences and at least one of the partial copies downstream *yhbQ*. This method was previously used to identify the pLP100 genomic island in *L. pneumophila* that can excise itself from the chromosome and exist as an episomal form (Trigui et al 2013).

The results of the PCR analysis (Figure 9) show multiple bands of different size which suggest that the region excises from the chromosome and forms an episome of different size containing *ybhQ* as well as a variable number of *lpr0024*. To confirm these results, the amplicons were cloned in pGEMT-easy vector and were sequenced by Sanger sequencing (Table 7). Total size of the amplified region 1116 bp, the size of each full copy of the sRNA is 158 bp so it is 632 bp for the four full. There is a partial copy on the left of *yhbQ* of 134 bp, on the right of *yhbQ* there are two partial copies of 90 bp and 72 bp. For instance, an amplicon of the size of 950 bp contains three full copies and the three partial copies, it is an episome because the recombination happened with the partial copies on the right of *yhbQ*.

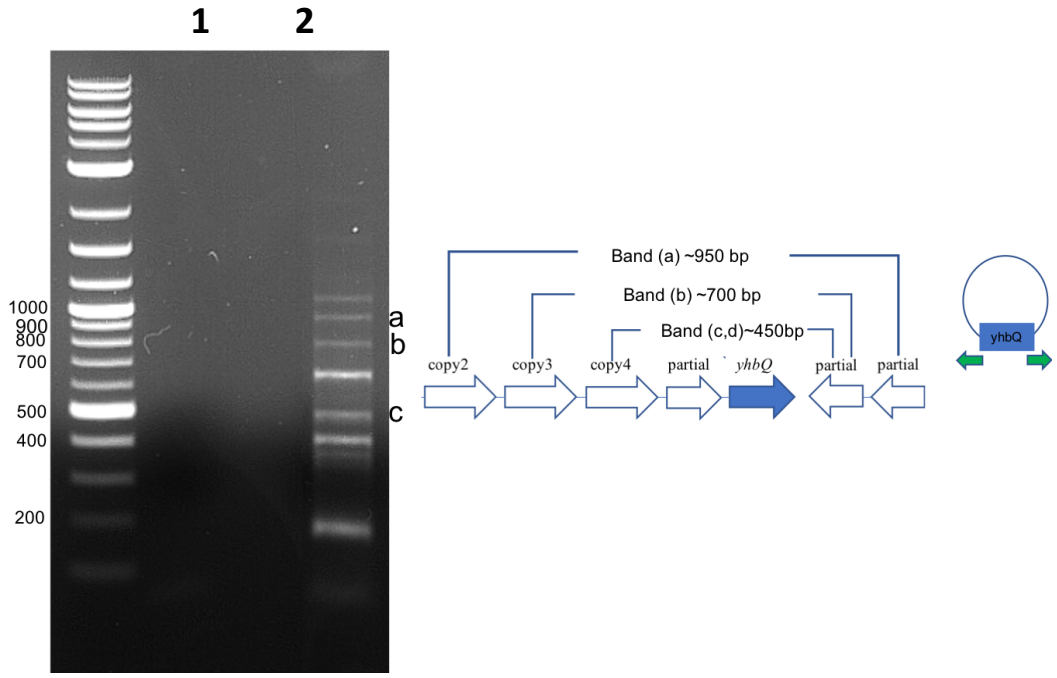

Figure 9. PCR analysis and sequencing to test the excision of Lpr0024 genomic region. Excision was tested by PCR using the primers episome-F and episome-R. Samples are: 1, negative control (water); 2, DNA from the wild-type (KS79). Amplicons were cloned with pGEMT-easy and were sequenced by Sanger sequencing to confirm the episome formation. The genomic organization of some of the bands is depicted on the right side.

Table 7. Sequencing results to confirm the episome formation

Colony #	Size on Gel	Size based on sequencing	# of full copies based on sequencing	# of partial copies based on sequencing
1	~950 bp	1013 bp	3	1 upstream *yhbQ* & 2 downstream *yhbQ*
2	~500 bp	551 bp	1	1 upstream *yhbQ* & 1 downstream *yhbQ*
3	~450 bp	532 bp	1	1 upstream *yhbQ* & 1 downstream *yhbQ*
4	~700 bp	710 bp	2	1 upstream *yhbQ* & 1 downstream *yhbQ*

5. Discussion

Legionella pneumophila is a facultative intracellular pathogen and it is the causative agent of Legionnaires' disease, an acute form of pneumonia, and Pontiac fever, milder flu-like illness (Cazalet et al 2004, Weissenmayer et al 2011). It is a water borne pathogen which can be found in most water systems including freshwater bodies, rivers, and lakes as well as in engineered water systems and man-made water distribution systems such as cooling towers (Oliva et al 2017, Mendis et al 2015). Free-living amoeba in aquatic environments are the primary reservoir of *L. pneumophila*. It can be transmitted to human by inhalation of contaminated aerosols. Upon entry into the human's lung, *L. pneumophila* infects and replicates inside alveolar macrophages (Trigui et al 2013, Swanson and Hammer 2000).

L. pneumophila possesses a biphasic life cycle, switching between a replicative phase (RP), during infection of host cells, and a transmissive phase (TP), when the resources inside the cells are depleted and transmission to a new host cell is needed (Molofsky and Swanson 2004). Efficient regulation of gene expression is critical for the switch between the two phases as well as for intracellular multiplication and virulence (Bruggemann et al 2006). Small Regulatory RNAs (sRNAs) are major regulators of *L. pneumophila* gene expression (Faucher and Shuman 2011). This study investigates the role of two sRNA that have not been previously characterized, Lpr0050 and Lpr0024 which were identified by transcriptome analysis by RNA sequencing of *L. pneumophila* Philadelphia-1(Weissenmayer et al 2011).

Lpr0050 is a cis-encoded sRNA which is encoded on the complementary strand of its mRNA target *sdeA* (Oliva et l 2017). We were able to confirm by northern blot the expression of lpr0050 in *L. pneumophila* wild-type in E and PE phases (Figure 2). The expression of the sRNA was also tested in two different regulatory mutants, Δ*cpxR* and Δ*letS.* CpxR is the response regulator protein of

the two-component system CpxRA which was shown to directly activate or repress the expression of 12 effector-encoding genes and several *icm/dot* genes (Altman and Segal 2008). On the other hand, LetS is the sensor protein of the two-component system LetAS which is a critical regulator for differentiation of *L. pneumophila* in response to starvation inside the host cells as well as in broth culture (Edward et al 2010). The results showed that *lpr0050* is expressed in both mutants in the E and PE phases indicating that *lpr0050* is not regulated by any of these regulators. That could be an indication that this sRNA is not involved in the transmission of *L. pneumophila* from the replicative to the transmissive phase. Testing the expression of *lpr0050* in other regulatory mutants maybe required to gain more knowledge on its potential role.

Lpr0050 is encoded on the complementary strand of the Icm/Dot effector SdeA (Bradill 2007). In order to study the function of a specific gene, one of the methods used is creating a deletion mutation of that gene. However, such a method cannot be used to study the role of *lpr0050* because it is a cis-encoded sRNA and it is complementary to its mRNA target which is *sdeA*. In order to study the role of *lpr0050*, a fragment of *sdeA* complementary to the sRNA was cloned under the *Ptac* promoter. This produces an RNA complementary to the sRNA, called the trap. Our goal was to deplete the sRNA without making changes in the mRNA target, *sdeA*, to identify its function. The rationale behind this method is that overexpression of the trap could affect the expression of lpr0050 or block its function by sequestering and therefore preventing it from interacting with its target, *sdeA*, thus relieving the effect of lpr0050. A 300 bp fragment of *sdeA* was overexpressed by cloning in pMMB207c (Table 4) under the *Ptac* promoter. In order to test if the trap was cloned successfully, we performed northern blot to detect its expression using the designed oligonucleotides probes (Table 5). It is shown in Figure 3 A and B, that the 300 bp trap was expressed in SPF482 (Table 1) grown with 1mM IPTG.

The results show that the trap is expressed in the E phase but not in PE phase, the reason of which is unknown. Overexpression of the trap had no effect on the sRNA expression or stability (Figure 3C , D). If the expression of *lpr0050* was reduced by the overexpression of the trap, that would be explained as sRNA-mRNA degradation (Waters and Storz, 2009). This would be similar to the IsrR/*isiA* system involved in photosynthesis in *Synechocystis* sp. PCC 6803 where IsrR-mediated *isiA*-mRNA degradation has been suggested (Duhring et al 2006).

On the other hand, overexpression of the trap may somehow affect the function of the sRNA which in turn affect the expression of the mRNA target. Cis-encoded sRNAs can control gene expression by binding to their mRNA targets on the complementary strand. The binding occurs at the 5′ end, 3′ end or in the middle of mRNAs, depending where it is encoded (Cho and Kim 2015). Based on their binding sites on their mRNA targets, cis-encoded sRNAs may act as transcription terminators through transcription attenuation or transcription interference, potential inhibitors of translation initiation, or modulators of mRNA degradation (Cho and kim 2015).

In transcription interference, two promoters of an antisense RNA and its target sense RNA exist very close in *cis*-position and their transcriptions occur in the convergent direction and then the transcription from one promoter becomes suppressed by the other one (Callen et al, 2004). Transcriptional interference can occur due to the collision of the RNA polymerases rather than sRNA binding to the mRNA (Prescott and Proudfoot 2002). This can be evaluated by using reporter genes which encode protein that can be easily assayed (Herschman 2004). Reporters are very efficient tools to evaluate the efficacy of gene expression (Gambhir 2002). Among the conventional reporter genes that have been commonly used are those that encode for the following proteins: chloramphenicol acetyltransferase, β-galactosidase (GAL), β-lactamase and green fluorescent protein (GFP) (Nomura et al 1997, Lee et al 1999, Sauvonnet and Pugsley1996,

Kunert et al 2000). The introduced reporter gene is driven by a promoter of choice (Gambhir 2002). The reporter gene is cloned downstream of a regulatory region such as promoter or enhancer that is usually responsible for the controlled expression of a specific gene (*sdeA* in this case). This would permit the evaluation of transcriptional interference due to the collision of the RNA polymerases (Gambhir 2002).

The expression of *sdeA*, the mRNA target of *lpr0050*, was measured in the E phase (Figure 4B) since the trap was expressed only in the E phase (Figure 3B). The expression of *sdeA* was measured before the binding site at the 5' and after the binding site at the 3' end to be able to test if the sRNA would affect the expression of the whole transcript or only the section downstream of the binding site. As the results reveal, there was no significant difference in the expression of *sdeA* at the 5'end whether the trap was expressed or not. However, at the 3' end the expression of *sdeA* was reduced when the trap was overexpressed. The results indicate that sequestering the sRNA lowers the expression of *sdeA* (significantly at the 3' end) which suggests that *lpr0050* might be necessary to maintain expression of *sdeA* along the whole transcript, maybe by increasing its stability as was shown for other sRNAs (Waters and Storz 2009).

The effect on SdeA expression could be merely some indirect effect of overexpression of an RNA. For instance, overexpression of the trap could titer down Hfq which in turn might affect the expression of SdeA indirectly. To test this, the expression of SdeA could be measured after overexpression of a "trap scrambled sequence". If overexpression of the scrambled trap has no effect on SdeA, it means that the changes of SdeA expression was due the trap overexpression not because it is just some indirect effect due to overexpression of an RNA.

The expression of *sdeA* homologs was also measured in the E phase using the same strategy as for *sdeA* (Figure 4 B). As the results show, there is significant difference in the expression of whole

transcript of *sdeA* homologs before and after the binding site upon overexpression of the trap. Thus, in contrast to *sdeA*, *lpr0050* could downregulate *sdeA* homologs maybe by inducing the decay of the target, or by inhibiting translation (Faucher and Shuman 2011, Waters and Storz 2009, Brantl 2007). Therefore, blocking the function of *lpr0050* by sequestration with the trap induces the expression of *sdeA* homologs. Taken together, these results suggest that *sdeA* and its homologs are co-regulated through *lpr0050*.

Inhibition of translation cannot be tested by qPCR because it is used to measure gene expression at the transcriptional level only. Therefore, other techniques such as western blot can be used to test the expression of SdeA and its homologs at the protein level which could confirm the results on regulation of SidE family by Lpr0050.

L. pneumophila is an excellent model organism to study the interactions between secreted effector proteins, as its Dot/Icm T4SS translocate around 300 substrates among which SidE effector proteins and SidJ (Jeong et al 2015). *sidJ* is located on the *sdeC-sdeA* locus, but it is not part of the SidE family (Liu and Luo 2007). It has been shown that *sidJ* functions as a metaeffector to regulate the activity of the SidE substrates (Jeong et al 2015). We have measured the expression of *sidJ* in a similar manner as for *sdeA* and its homologs (Figure 4B). As the results show, when the trap is overexpressed there is lower expression of *sidJ* (before and after the binding site), indicating that Lpr0050 is a positive regulator of SidJ.

As mentioned, Lpr0050 represses all the SidE homologous but induces the expression of SdeA and SidJ. On the other hand, the gradual accumulation of SidJ during the infection and the decreased level of the SidE family on LCV are corelated (Jeong et al 2015). Furthermore, SdeA-toxicity could be suppressed by co-expression of SidJ, implying that SidJ may regulate SdeA function. However, secretion of SdeA was not affected by the absence of *sidJ* (Jeong et al 2015).

All together suggests that *lpr0050* is important in SidJ regulation and that both *lpr0050* and SidJ repress SidE homologues. Additionally, *lpr0050* is important for the expression of the whole transcript of SdeA whose secretion is not affected by *sidJ* (Jeong et al 2015).

Overall, SdeA and its homologs are substrates of Dot/Icm T4SS, and they were found to be important in the intracellular multiplication of *L. pneumophila* (Jeong et al 2015). The results show that *lpr0050* is involved in the regulation of *sdeA* and its homologs, therefore lpr0050 could play a role in the intracellular multiplication of *L. pneumophila.* Further experiments and analysis is needed to confirm the role of lpr0050 in *L. pneumophila* virulence. Testing the expression of SdeA and its homologs at the protein level could confirm the results on regulation of SidE family by *lpr0050*. Moreover, infection of host cells, cell line THP-1, using SPF482 (ON and OFF) to examine the effect of the sRNA in the intracellular multiplication and virulence of *L. pneumophila.*

Lpr0024 on the other hand represents an example of trans-encoded sRNAs which are usually encoded in intergenic regions on the chromosome and show only partial complementary to their target mRNAs (Cho and kim 2015).

Expression of *lpr0024* in the wild-type was confirmed by northern blot (Figure 6B). These results indicate that *lpr0024* is regulated by the LetAS TCS in the PE phase. In *L. pneumophila* LetAS TCS is responsible for differentiation in response to starvation (Mendis et al 2018). Therefore, it could be suggested that lpr0024 is involved for differentiation of *L. pneumophila* and for the transmission from the replicative phase to the transmissive phase (Edwards et al 2010).

Genetic analysis of *lpr0024* is represented in (Figure 7). The sRNA is located between *iraB* and *lpg0745* genes. There is also an unannotated ORF with a putative endonuclease function present in the genomic region. We called the ORF *yhbQ* because it is homologous to the predicted endonuclease, GIY-YIG superfamily, in *E. coli* (Kumar et al 2016). It was noted that *lpr0024* is a

multicopy sRNA, four full copies of the sRNA are present in the Philadelphia-1 genome and there are three partial copies, one is encoded upstream of *yhbQ* whereas the other two are encoded downstream of it. Thus, *lpr0024* represents a multicopy sRNA in *L. pneumophila.*

Multicopy sRNAs have been previously studied in the Gram positive bacteria *Listeria monocytogenes* (Sievers, et al 2014). The multicopy sRNA LhrC contains five homologous sRNAs, called LhrC1-5 (Sievers, et al 2015). These five nearly identical sRNAs are highly induced in response to cell envelope stress and target the virulence adhesin *lapB* at the post-transcriptional level (Mollerup et al 2016). They control gene expression by base pairing to target mRNAs though three conserved UCCC motifs common to all 5 LhrCs. (Sievers, et al 2015). multicopy sRNAs can be an important feature of bacterial gene regulation (Sievers, et al 2015).

Because the lpr0024 sequence was found repeated multiple times, we hypothesized that recombination events may occur between them which might lead to variation in the copy numbers of the sRNA at the population level. Homologous recombination (HR) generates genome rearrangements involving repeated DNA elements with identical (homologous) or near-identical sequences that can be located anywhere in the genome (Piazza and Heyer2018). In order to confirm if there are recombination events occurring between the repeated sequences of *lpr0024*, we performed PCR analysis of the region between *iraB* and *yhbQ* and sequencing of the resulting amplicons (Figure 8). The PCR results revealed multiple bands of different size ranging from 500 bp to 1200 bp. The sequencing results confirms that the number of copies varies at the population level between 1 to 5 copies. There are a number of explanations about the benefits of homologous recombination in bacteria (David et al 2017). One explanation is that recombination is used as a repair mechanism of DNA damage using foreign DNA as a template, this is essential to maintain genome integrity (Michod et al 2008, Piazza and Heyer2018). Another explanation is that

recombination events are able to remove deleterious mutations and introduce combinations of advantageous mutations which could increase the efficiency of natural selection (Narra et al 2006). Finally, a recent study has also suggested that bacteria use recombination to delete selfish mobile genetic elements from their genomes (Croucher et al 2008). Previous studies have shown that recombination accounts for more than 96% of diversity within several major disease-associated sequence types (STs) of *L. pneumophila*. This suggests that recombination represents a potentially important player for adaptation and virulence (David et al 2017). Therefore, it can be suggested that lpr0024 plays a role in *L. pneumophila* adaptation and virulence since recombination events occur between its repeated copies.

Recombination between one of the multiple copies and one of the inverted partial copies encoded after *yhbQ* could result in the excision of this region and formation of an episome. We tested this by PCR analysis and Sanger sequencing (Figure 9). The PCR results show that there are multiple bands of different size (Figure 8) which indicates there is formation of an episome containing, at the population level, different copies of *lpr0024*. Additionally, the sequencing results (Table 7) revealed that the recombination is happening between at least one full copy and one partial copy encoded downstream of *yhbQ*, indicating that this region seems to be a mobile element able to excises itself from the genome and exists in an episomal form. Figure 10 below illustrates the excision of the genomic region and the episome formation. The episome contains several copies of the sRNA as well as the endonuclease *yhbQ.* Thus, analysis of lpr0024 suggests this is a new mobile genetic element (MGE) which would presumably use YhbQ to excise itself.

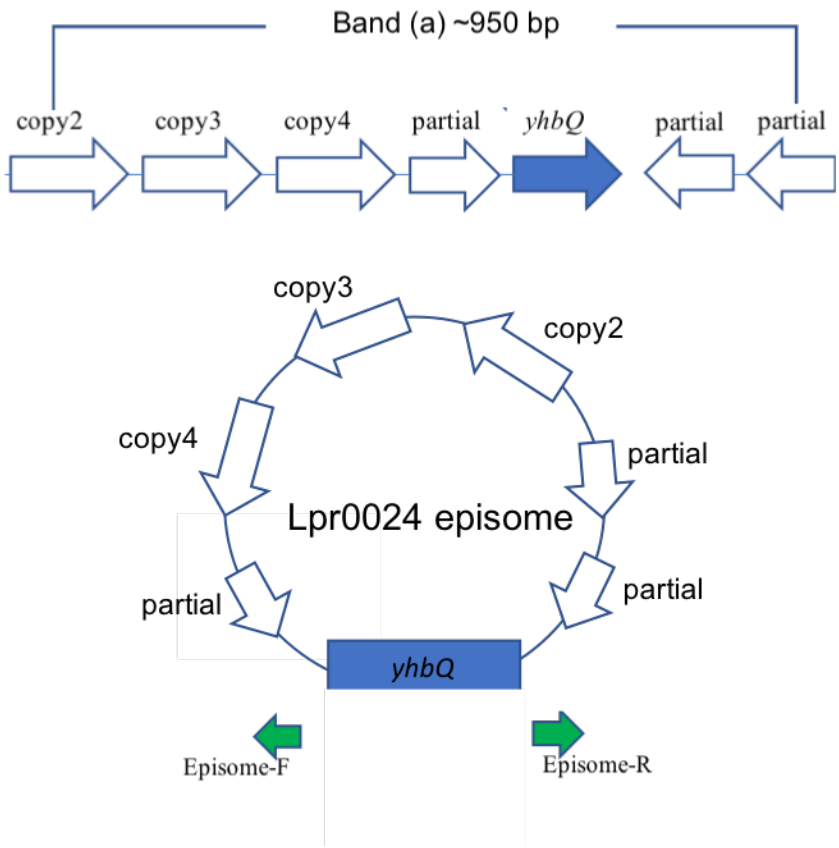

Figure 10. Schematic representation of the excision of lpr0024 genomic region in the chromosome of *L. pneumophila*. The primers used in this study are episome-F and episome-R (in green arrows). Four clones were sequenced to confirm the episome formation, the represented example is for a clone of ~950 bp which contains three full copies, one partial copy before the *yhbQ*, and the two partial copies after *yhbQ*

It has been shown that MGE are major players in genome plasticity of *L. pneumophila* (Dagan et al 2008). Trigui et al (2013) have identified a genomic island of 100 kb representing a mobile genetic element (Trigui et al 2013 b). Furthermore, another study showed that the pLP45 possesses a sRNA (lpr0035) encoded in the *attL* site that regulates a few virulence-related genes encoded within it (Jayakumar et al 2012).

Hfq is essential in preventing excision of pLP100 from the chromosome during the PE phase (Trigui et al 2013 b). Therefore, further experiments are needed to investigate how excision of lpr0024 episome is regulated and what is the regulation mechanism. The amplicon that has one

copy could be cloned into a plasmid and transferred to the wild-type. To confirm the episome insertion, primers for that plasmid can be designed and by PCR analysis we can test if the episome moved into the plasmid.

In the Paris strain, excision of pP36, a 36-kbp mobile genetic element containing an Lvh T4ASS locus, from the chromosome was observed only during the PE phase (Dole´ans-Jordheim et al 2006), which suggests that transcriptional programs active during PE might have an effect on the excision of integrated elements in *L. pneumophila* (Trigui et al 2013 b). Potentially, stress that *L. pneumophila* encounters during PE phase, such as starvation or drop in pH could affect excision of these elements from the chromosome (Trigui et al 2013 b). Thus, formation of lpr0024 episome could be tested under various stress conditions during PE phase to understand the conditions and the mechanism of the episome formation

Glockner et al reported that the mobile genetic element that contains Trb-1 T4ASS exhibited increased excision in *L. pneumophila* Corby after intracellular multiplication in *A. castellanii* (Glockner et al 2008), supporting a correlation between mobility of the element and virulence phenotypes (Jayakumar et al 2012). Furthermore, the excision of pLP100 results in induction of the *copA* gene, which increases resistance of *L. pneumophila* to copper (Trigui et al 2013 b). Thus, additional investigation is required to determine what are the phenotypes linked to lpr0024 episome.

In order to study the role of lpr0024 in the virulence of *L. pneumophila*, deleting the whole region of the repeated sequence to create a mutant is required. We tried to delete the whole region but it was unsuccessful. Therefore, an alternative method should be used such as direct transformation of *E. coli* carrying a Red helper plasmid with PCR products having short homology extensions for the targeted locus. This method was used to isolated chromosomal mutants with 13 different gene

disruptions (Datsenko And Wannar 2000). This method is also used for efficient generation of deletion mutants in *Legionella pneumophila* (Bryan et al 2011)**.** The basic strategy is to replace a chromosomal sequence (lpr0024 gene in this case) with a selectable antibiotic resistance gene that is generated by PCR by using primers with homology extensions. This could be done by Red-mediated recombination in these flanking homologous sequences. After selection, the resistance gene can also be eliminated by using a helper plasmid expressing the FLP recombinase, which acts on the directly repeated FRT (FLP recognition target) sites flanking the resistance gene (Datsenko And Wannar 2000). This method was also used to delete *moaA* gene in *S. typhimurium* (Koskiniemi et al 2012).

Enhancing virulence and intracellular multiplication are among the predicted phenotype as well as adaptation to hard conditions or stress. To confirm the phenotype, a complemented strain could be constructed by complementing the clone that has only one copy into a plasmid and test if it can restore the WT phenotype. The benefit of having the multicopy region can be also investigated by making complemented strains with one copy or the whole region to test if one copy is enough for any given phenotype or the presence of the whole region is required.

6. Conclusions

This study sheds lights on two uncharacterized small regulatory RNAs that were previously discovered through the transcriptome sequencing of *L. pneumophila* Philadelphia-1. The results suggest that the cis-encoded sRNA Lpr0050 is expressed in the E and PE phases. Our results indicate that *sdeA* and its homologos*, sdeC, sdeB and sidE,* are co-regulated through *lpr0050.* Additionally, *sidJ,* metaeffector of SidE family, is also regulated by lpr0050. Further analysis and investigation is required to investigate the mode of action of lpr0050 and its role in the regulation of virulence *of L. pneumophila.*

On the other hand, we showed that the trans-encoded Lpr0024 is expressed in the E and PE phases in the wild-type and it is regulated by the TCSs CpxRA and LetAS in the PE phase. PCR and sequencing analysis revealed that this sRNA is a multicopy sRNA and that the number of copies varies between 1 to 5 copies at the population level, as a result of active recombination between the copies. Lastly, we were able to show that the region of lpr0024 can excise from the genome and exist in an episome.

Further work is required to investigate the regulatory role of Lpr0024. In order to do that, a mutant should be constructed by deleting the whole region of lpr0024. This was tried in this study (data not shown) by creating allelic exchange fragments containing kanamycin resistance cassette by using long flanking homology PCR but it was not successful. In this study we tried to generate a mutant strain for *lpr0024* by creating a mutant allele containing a kanamycin resistance cassette using the long-flanking homology PCR technique, but we were unsuccessful. Therefore, an alternative method should be used such as direct transformation of *E. coli* carrying a Red helper plasmid with PCR products having short homology extensions for the targeted locus.

What is the benefit of having the sRNA in multiple copies? And are all the copies required for full virulence? This can be investigated by making complemented strains with one copy or the whole region to test if one copy is enough for any given phenotype or the presence of the whole region is required. These are questions that need to be addressed by future work on this sRNA.

7. References

- Abu Kwaik Y. (1996). The phagosome containing Legionella pneumophila within the protozoan Hartmannella vermiformis is surrounded by the rough endoplasmic reticulum. *Appl Environ Microbiol* 62(6):2022–2028
- Abu-Zant A, Asare R, Graham J, Abu Kwaik Y (2006). Role for RpoS but not RelA of Legionella pneumophila in modulation of phago-some biogenesis and adaptation to the phagosomal microenvironment. *Infect Immun* 74:3021–3026
- Al-Quadan T, Price CT, Kwaik YA (2012). Exploitation of evolutionarily conserved amoeba and mammalian processes by Legionella. *Trends Microbiol* 20(6):299–306
- Altman E, Segal G (2008). The response regulator CpxR directly regulates expression of several Legionella pneumophila icm/dot components as well as new translocated substrates. *J Bacteriol* 190:1985–1996
- Amaro F., Gilbert J. A., Owens S., Trimble W. Shuman H. A. (2012). Whole-genome sequence of the human pathogen Legionella pneumophila serogroup 12 strain 570-CO-H. *J Bacteriol* 194, 1613– 1614.
- Amaro F., Wang W., Gilbert J. A., Anderson O. R., Shuman, H. A. (2015). Diverse protist grazers select for virulence-related traits in Legionella. ISME J. 9, 1607–1618. doi: 10.1038/ismej.2014.248
- Amor, J. C., J. Swails, X. Zhu, C. R. Roy, H. Nagai, A. Ingmundson, X. Cheng, and R. A. Kahn. (2005). The structure of RalF, an ADP-ribosylation factor guanine nucleotide exchange factor from Legionella pneumophila, reveals the presence of a cap over the active site. *J. Biol. Chem*. 280:1392– 1400
- Babitzke P., Romeo T. (2007). CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol*. 10:156–163.
- Backert S., Grohmann E. (2018). Type IV Secretion in Gram-Negative and Gram-Positive Bacteria. *Curr Top Microbiol Immunol*. 413
- Bardill J. P., Miller J.L, Voge J.P. (2005). IcmS-dependent translocation of SdeA into macrophages by the *Legionella pneumophila* type IV secretion system. *Mol. Microbiol*. 56:90–103.
- Bartley PB, Ben Zakour NL, Stanton-Cook M, et al. (2016). Hospital-wide eradication of a nosocomial *Legionella pneumophila* serogroup 1 outbreak. *Clin Infect Dis*. 62:273–279.

- Beer KD, Gargano JW, Roberts VA, et al (2015). Surveillance for waterborne disease outbreaks associated with drinking water - United States, 2013–2014. *MMWR Morb Mortal Wkly Rep*. 64:842–848.
- Bellinger-Kawahara C, Horwitz MA (1990). Complement component C3 fixes selectively to the major outer membrane protein (MOMP) of *Legionella pneumophila* and mediates phagocytosis of liposome-MOMP complexes by human monocytes. *J Exp Med* 172(4):1201–1210
- Berger, K. H., J. J. Merriam, and R. R. Isberg. (1994). Altered intracellular targeting properties associated with mutations in the Legionella pneumophila dotA gene. *Mol. Microbiol*. 14:809–822
- Benson RF, Fields BS. Classification of the genus *Legionella* (1998). Semin Respiratory Infect. 13(2):90–9.
- Bohn, C., Rigoulay, C., and Bouloc, P. (2007). No detectable effect of RNA binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.* 7, 10. doi: 10.1186/1471-2180-7-10
- Brantl S (2002). Antisense-RNA regulation and RNA interference. *Biochim Biophys Acta*. 1575:15-25.
- Brassinga AK, Hiltz MF, Sisson GR, Morash MG, Hill N, Garduno E, Edelstein PH, Garduno RA, Hoffman PS (2003). A 65-Kilobase Pathogenicity Island Is Unique to Philadelphia-1 Strains of Legionella pneumophila. *J Bacteriol* . 185(15):4630-3637.
- Brenner, D. J., A. G. Steigerwalt, and J. E. McDade. (1979). Classification of the Legionnaires' disease bacterium: *Legionella pneumophila*, genus novum, species nova, of the family Legionellaceae, familia nova. *Ann. Intern. Med*. 90:656–658.
- Bruggemann H, Hagman A, Jules M, Sismeiro O, Dillies MA, Gouyette C, Kunst F, Steinert M, Heuner K, Coppee JY *et al* (2006).Virulence strategies for infecting phagocytes deduced from the in vivo transcriptional program of *Legionella pneumophila*. *Cell Microbiol* 8(8):1228-1240.
- Bryan A, Harada K, Swanson MS. (2011). Efficient Generation of Unmarked Deletions in *Legionella pneumophila. Appl. Environ. Microbiol.* 77(7): 2545-2548
- Buchrieser C, Hilbi H (2019). *Legionella* Methods and Protocols. https://doi.org/10.1007/978-1-4939-9048-1

- Byrne, B., and Swanson, M.S. (1998). Expression of Legionella pneumophila virulence traits in response to growth conditions. *Infect Immun* 66: 3029–3034
- Callen BP, Shearwin KE, and Egan JB. (2004). Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol. Cell* 14, 647–656
- Campodonico, E. M., L. Chesnel, and C. R. Roy. (2005). A yeast genetic system for the identification and characterization of substrate proteins transferred into host cells by the Legionella pneumophila Dot/Icm system. *Mol. Microbiol*. 56:918–933.
- Cavanagh AT, Klocko AD, Liu X, Wassarman KM (2008). Promoter specificity for 6S RNA regulation of transcription is determined by core promoter sequences and competition for region 4.2 of σ70. *Mol Microbiol* 67:1242–1256.
- Cho KH and Kim J-H (2015) Cis-encoded non-coding antisense RNAs in streptococci and other low GC Gram (+) bacterial pathogens. *Front. Genet*. 6:110. doi: 10.3389/fgene.2015.00110
- Cazalet C et al (2004). Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity. *Nat Genet* 36:1165–1173.
- Cazalet C, et al. Analysis of the *Legionella longbeachae* Genome and Transcriptome Uncovers Unique Strategies to Cause Legionnaires' Disease. *PLoS Genetics* 2010: 6(2):1-16
- Chao, Y, and Vogel, J. (2010). The role of Hfq in bacterial pathogens *Curr. Opin. Microbiol.* 13, 24-33
- Chaudhry R, Sreenath K, Agrawal SK, Valavane A. (2018). *Legionella a*nd Legionnaires' disease: Time to explore in India. *Indian J Med Microbiol.* 36(3):324-333.
- Chen, D.-Q., Huang, S.-S. & Lu, Y.-J. (2006). Efficient transformation of *Legionella pneumophila* by high-voltage electroporation. *Microbiol Res* 161, 246–251.
- Chien M, Morozova I, et al (2004). The genomic sequence of the accidental pathogen *Legionella pneumophila. Science* 305:1966–1968.
- Cianciotto NP (2001). Pathogenicity of *Legionella pneumophila*. *Int J Med Microbiol* 291: 331-343.
- Cianciotto, N. P. (2007). Iron acquisition by *Legionella pneumophila*. *Biometals* 20, 323-331

- Cirillo JD, et al. (1999). Intracellular growth in *Acanthamoeba castellanii* affects monocyte entry mechanisms and enhances virulence of *Legionella pneumophila*. *Infect. Immun*. 67:4427– 4434.
- Cirillo JD, Falkow S, Tompkins LS. (1994). Growth of *Legionella pneumophila* in *Acanthamoeba castellanii* enhances invasion. *Infect. Immun*.62:3254 –3261.
- Cirillo SL, Bermudez LE, El-Etr SH, Duhamel GE, Cirillo JD (2001). *Legionella pneumophila* entry gene rtxA is involved in virulence. *Infect Immun* 69: 508-517
- Cirillo SL, Lum J, Cirillo JD (2000). Identification of novel loci involved in entry by *Legionella pneumophila. Microbiology* 146: 1345-1359
- Coil DA, Anné J (2009). Twitching motility in *Legionella pneumophila*. *FEMS Microbiol Lett* 293: 271-277.
- Collier SA, Stockman LJ, Hicks LA, et al. (2012). Direct healthcare costs of selected diseases primarily or partially transmitted by water. *Epidemiol Infect.* 140:2003–2013
- Conover, G. M., Martinez-Morales, F., Heidtman, M. I., Luo, Z. Q., Tang, M., Chen, C., Geiger, O., and Isberg, R. R. (2008). Phosphatidylcholine synthesis is required for optimal function of *Legionella pneumophila* virulence determinants *Cell. Microbiol* 10, 514-528
- Correia, A. M., Ferreira, J. S., Borges, V., Nunes, A., Gomes, B., Capucho, R., et al. (2016). Probable person-to-person transmission of Legionnaires' disease. *N. Engl. J. Med*. 374, 497–498.
- Croucher NJ, Mostowy R, Wymant C, Turner P, Bentley SD, Fraser C.
- 92016) Horizontal DNA Transfer Mechanisms of Bacteria as Weapons of Intragenomic Conflict. *PLoS Biology*. 14(3).
- D'Auria, G., Jime´ nez-Herna´ ndez, N., Peris-Bondia, F., Moya, A. & Latorre, A. (2010). *Legionella pneumophila* pangenome reveals strain specific virulence factors. *BMC Genomics* 11, 181
- Derre I, Isberg RR (2005). LidA, a translocated substrate of the *Legionella pneumophila* type IV secretion system, interferes with the early secretory pathway. *Infect Immun* 73:4370–4380.
- Dagan, T., Artzy-Randrup, Y. & Martin, W. (2008). Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc Natl Acad Sci U S A* 105, 10039–10044.

- Dalebroux, Z. D., Edwards, R. L. & Swanson, M.S. (2009). SpoT governs *Legionella pneumophila* differentiation in host macrophages. *Mol Microbiol* 71, 640–658.
- Datsenko KA, Wanner BL (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640–6645.
- David S, SaÂnchez-BusoÂ L, Harris SR, Marttinen P, Rusniok C, Buchrieser C, et al. (2017). Dynamics and impact of homologous recombination on the evolution of *Legionella pneumophila*. *PLoS Genet 13*(6): e1006855. https://doi.org/10.1371/journal.pgen.1006855
- de Felipe KS, Pampou S, Jovanovic OS, Pericone CD, Ye SF, Kalachikov S, Shuman HA (2005). Evidence for acquisition of *Legionella* type IV secretion substrates via interdomain horizontal gene transfer. *J Bacteriol* 187:7716–7726.
- Derre, I., and R. R. Isberg. (2004). *Legionella pneumophila* replication vacuole formation involves rapid recruitment of proteins of the early secretory system. *Infect. Immun*. 72:3048–3053
- Doleans-Jordheim A, et al. (2006). Growth-phase-dependent mobility of the lvh-encoding region in *Legionella pneumophila* strain Paris. *Microbiology* 152:3561–3568.
- Doleans, A., Akermi, M., Ginevra, C., Cazalet, C., Kay, E., Schneider, D., Buchrieser, C., Atlan, D., Vandenesch, F., Etienne, J. and Jarraud, S. (2006). Growth phase-dependent mobility of the lvh- encoding region in *Legionella* strain. Paris. *Microbiology* 152, 3561–3568.
- Doleans, A., Aurell, H., Reyrolle, M., Lina, G., Freney, J., Vandenesch, F., Etienne, J. and Jarraud, S. (2004). Clinical and environmental distributions of *Legionella* strains in France are different. *J. Clin. Microbiol*. 42, 458–460.
- Dooling KL, Toews KA, Hicks LA, et al (2015). Active bacterial core surveillance for legionellosis -United States, 2011–2013. *MMWR Morb Mortal Wkly Rep*. 64:1190–1193.
- Duhring U, Axmann IM, Hess WR, Wilde A (2006). An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc Natl Acad Sci USA*. 103:7054-7058.
- Dunin-Horkawicz S, Feder M, Bujnicki JM (2006). Phylogenomic analysis of the GIY-YIG nuclease superfamily. *BMC Genomics.* 7:98.
- Dupré, E. et al (2015). Virulence Regulation with Venus Flytrap Domains: Structure and Function of the Periplasmic Moiety of the Sensor-Kinase BvgS. *PLo Pathog* 11, e1004700, https://doi.org/10.1371/journal.ppat.1004700.

- Edwards RL, Jules M, Sahr T, Buchrieser C, Swanson MS. (2010). The *Legionella pneumophila* LetA/LetS two-component system exhibits rheostat-like behavior. *Infect Immun*. 78(6):2571-83. doi: 10.1128/IAI.01107-09.
- Eisenreich W, Dandekar T, Heesemann J, Goebel W. (2010). Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat Rev Microbiol* 8:401–412
- Ellis, T. N., and Kuehn, M. J. (2010). Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiol. Mol. Biol. Rev.* 74, *Rev.* 74, 81-94.
- Escoll P, Mondino S, Rolando M, Buchrieser C (2016). Targeting of host organelles by pathogenic bacteria: a sophisticated subversion strategy. *Nat Rev Microbiol* 14:5–19. https://doi.org/10. 1038/nrmicro.
- Escoll, P., Rolando, M., Gomez-Valero, L., and Buchrieser, C. (2013). From amoeba to macrophages: exploring the molecular mechanisms of *Legionella pneumophila* infection in both hosts. *Curr. Top.Microbiol. Immunol*. 376, 1–34.
- Faucher SP, Shuman HA (2011). Small Regulatory RNA and *Legionella pneumophila*. *Front Microbiol* 2:98.
- Faucher SP, Friedlander G, Livny J, Margalit H, Shuman HA (2010). *Legionella pneumophila* 6S RNA optimizes intracellular multiplication. *Proc Natl Acad Sci U S A* .107(16):7533-7538
- Feeley, J. C. et al (1979). Charcoal-yeast extract agar: primary isolation medium for *Legionella pneumophila*. *J Clin Microbiol* 10, 437–44.
- Fettes, P. S., Forsbach-Birk, V., Lynch, D., and Marre, R. (2001). Over expression of a *Legionella pneumophila* homologue of the *E. coli* regulator csrA affects cell size, flagellation, and pigmentation. *Int J. Med. Microbiol 291 353-360*
- Finsel I, Hilbi H (2015). Formation of a pathogen vacuole according to Legionella pneumophila: how to kill one bird with many stones. *Cell Microbiol* 17:935–950.
- Franze de Fernandez MT, Eoyang L, August JT. (1968). Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* 219: 588–590.
- Francione L, Smith PK, Accari SL, Taylor PE, Bokko PB, Bozzaro S et al (2009). *Legionella pneumophila* multiplication is enhanced by chronic AMPK signalling in mitochondrially diseased Dictyostelium cells. *Dis Model Mech* 2(9–10):479–489

- Fraser DW, Tsai TR, Orenstein W et al (1977) Legionnaires' disease. *New Engl J Med* 297:1189–1197
- Fraser, D. W., T. R. Tsai, W. Orenstein, W. E. Parkin, H. J. Beecham, R. G. Sharrar, J. Harris, G. F. Mallison, S. M. Martin, J. E. McDade, C. C. Shepard, and P. S. Brachman. (1977). Legionnaires' disease: description of an epidemic of pneumonia. *N. Engl. J. Med*. 297:1189–1197
- Florian Eggenhofer, Hakim Tafer, Peter F. Stadler, Ivo L. Hofacker (2011). RNApredator: fast accessibility based prediction of sRNA targets, *Nucleic Acids Research*. 39(2) W149–W154, https://doi.org/10.1093/nar/gkr467.
- Gal-Mor O, Segal G (2003a) Identification of CpxR as a positive regulator of *icm* and dot virulence genes of *Legionella pneumophila. J Bacteriol* 185:4908–4919
- Galka, F., Wai, S. N., Kusch, H., Engelmann, S., Hecker, M., Schmeck, B., Hippenstiel, S., Uhlin, B. E., and Steinert, M. (2008). Proteomic characterization of the whole secretome of *Legionella pneumophila* and functional analysis of outer membrane vesicles. *Infect. Immun.* 76, 1825-1836
- Gambhir SS (2002). Imaging of Reporter Gene Expression. *Brain Mapping The Methods* 2nd edition 799-818
- Garduño RA., Garduño E. Hoffman P. S. (1998). Surface -associated. hsp60 chaperonin of *Legionella pneumophila* mediates invasion in a HeLa cell model. *Infect. Immun.* 66,
- Glick TH, Gregg MB, Berman B, Mallison G, Rhodes WW, Kassanoff I (1978). Pontiac fever. *Amer J Epidemiol*. 107:149–160
- Glockner, G., Albert-Weissenberger, C., Weinmann, E., Jacobi, S., Schunder, E., Steinert, M., Hacker, J. & Heuner, K. (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, 411–428.
- Gomez-Valero, L., C. Rusniok, and C. Buchrieser. (2009). *Legionella pneumophila*: population genetics, phylogeny and genomics. *Infect. Genet. Evol*. 9:727–739.
- Gomez-Valero, L., Rusniok, C., Jarraud, S., Vacherie, B., Rouy, Z., Barbe, V., Medigue, C., Etienne, J. & Buchrieser, C. (2011). Extensive recombination events and horizontal gene transfer shaped the *Legionella pneumophila* genomes. *BMC Genomics* 12, 536.

- Gotoh Y, Eguchi Y, Watanabe T, Okamoto S, Doi A, Utsumi R (2010). Two-component signal transduction as potential drug targets in pathogenic bacteria. *Curr Opin Microbiol* 13:232–239
- Greub, G. and Raoult, D. (2004). Microorganisms resistant to free-living amoebae. *Clin Microbiol Rev* 17, 413–433.
- Grohmann E, Christie PJ, Waksman G, Backert S (2018). Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol Microbiol* 107:455–471.
- Gunn JS (2008). The Salmonella PmrAB regulon: lipopolysaccharide modifications, antimicrobial peptide resistance and more. *Trends Microbiol* 16:284–290
- Hales, L. M., and H. A. Shuman. (1999). *Legionella pneumophila* contains a type II general secretion pathway required for growth in amoebae as well as for secretion of the Msp protease. *Infect. Immun*. 67:3662–3666.
- Hales, L. M., and H. A. Shuman. (1999). The *Legionella pneumophila rpoS* gene is required for growth within *Acanthamoeba castellanii. J. Bacteriol*. 181:4879–4889.
- Hamilton KA, Prussin AJ, Ahmed W, Haas CN (2018). Outbreaks of Legionnaires' Disease and Pontiac Fever 2006-2017. *Curr Environ Health Rep.* 5(2):263-271.
- Hammer B, Tateda E, Swanson M (2002). A two-component regulator induces the transmission phenotype of stationary-phase *Legionella pneumophila. Mol Microbiol* 44:107–11
- Hardiman, C. A., McDonough, J. A., Newton, H. J. & Roy, C. R (2012). The role of Rab GTPases in the transport of vacuoles containing *Legionella pneumophila* and *Coxiella burnetii*. *Biochem. Soc. Trans.* 40, 1353–1359.
- Helbig, J. H., and Amemura-Maekawa, J. (2009). Serotyping of *Legionella pneumophila* and epidemiological investigations. *Nowa Medycyna* 16, 69-75.
- Hengge-Aronis R (2002). Signal Transduction and Regulatory Mechanisms Involved in Control of the S (RpoS) Subunit of RNA Polymerase. *Microbiology and Molecular Biology Reviews.* 66(3):373-395
- Hershko A, Ciechanover A, Varshavsky A (2000). Basic Medical Research Award. The ubiquitin system. *Nat Med* .6:1073-1081.
- Herschman HR (2004). Noninvasive Imaging of Reporter Gene Expression in Living Subjects. *Advances in cancer research*. 92: 29-80

- Herwaldt L.A, Marra A. R (2018). Legionella: a reemerging pathogen *Curr Opin Infect Dis.* ;31(4):325-333.
- Heuner, K., and Albert-Weissenberger, C. (2008). *Legionella Molecular Biology*. Norfolk: *Caister Academic Press.*
- Hilbi H, Hoffmann C, Harrison CF (2011). Legionella spp. outdoors: colonization, communication and persistence. *Environ Microbiol Rep.* 3(3):286-96.
- Hilbi H, Molecular Mechanisms in *Legionella* Pathogenesis (2014), DOI 10.1007/978-3-642-40591-4
- Hilbi H, Segal G, Shuman HA (2001). Icm/dot-dependent upregulation of phagocytosis by *Legionella pneumophila. Mol Microbiol* 42: 603- 617.
- Hilbi, H.,Weber, S. S., Ragaz, C.,Nyfeler, Y., and Urwyler, S. (2007). Environmental predators as models for bacterial pathogenesis. *Environ. Microbiol*. 9, 563–575.
- Hilbi H, Nagai H, Kubori T, Roy CR (2017) Subversion of Host Membrane Dynamics by the *Legionella* Dot/Icm Type IV Secretion System. *Curr. Top. Microbiol. Immunol.* 413: 221–242
- Hindley, J. (1967). Fractionation of 32P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. *J. Mol. Biol.* 30, 125-136.
- Hoffmann, C., Finsel, I., and Hilbi, H. (2013). Pathogen vacuole purification from *Legionella*-infected amoeba and macrophages. *MethodsMol. Biol.* 954, 309–321.
- Horwitz, M. A. (1983). Formation of a novel phagosome by the Legionnaires' disease bacterium (*Legionella pneumophila*) in human monocytes. *J. Exp. Med*. 158:1319–1331.
- Horwitz, M. A., Maxfield, F. R (1984). *Legionella pneumophila* inhibits acidification of its phagosome in human monocytes. *J. Cell Biol*. 99, 1936–1943.
- Hovel-Miner G, Pampou S, Faucher SP, Clarke M, Morozova I, Morozov P, Russo JJ, Shuman HA, Kalachikov S (2009). SigmaS controls multiple pathways associated with intracellular multiplication of *Legionella pneumophila. J Bacteriol* . 191(8):2461-2473.
- Hubber A, Roy CR (2010). Modulation of host cell function by *Legionella pneumophila* type IV effectors. *Annu Rev Cell Dev Biol* 26:261–283.
- Hung, D. L., T. L. Raivio, C. H. Jones, T. J. Silhavy, and S. J. Hultgren. (2001). Cpx signaling pathway monitors biogenesis and affects assembly and expression of P pili. *EMBO J.* 20:1508–1518

- Huws, S.A., McBain, A.J. and Gilbert, P. (2005). Protozoan grazing and its impact upon population dynamics in biofilm communities. *J Appl Microbiol* 98, 238–244.
- Iriarte, M., I. Stainier, and G. R. (1995). The *rpoS* gene from *Yersinia enterocolitica* and its influence on expression of virulence factors. *Infect. Immun*. 63:1840–1847.
- Isberg RR, O'Connor TJ, Heidtman M (2009) .*The Legionella pneumophila* replication vacuole: making a cosy niche inside host cells. *Nat Rev Microbiol* 7(1):13–24
- Jayakumar D, Early JV, Steinman HM. (2012). Virulence phenotypes of *Legionella pneumophila* associated with noncoding RNA lpr003 *Infect Immun.* 80(12):4143-53.
- Jules M, Buchrieser C (2007). *Legionella pneumophila* adaptation to intracellular life and the host response: clues from genomics and transcriptomics. *FEBS Lett.* 581(15):2829-2838.
- Jung K, Fried L, Behr S, Heermann R (2012). Histidine kinases and response regulators in networks. *Curr Opin Microbiol* 15:118–124.
- Kagan JC, Roy CR (2002). *Legionella phagosomes* intercept vesicular traffic from endoplasmic reticulum exit sites. *Nat Cell Biol* 4:945–954.
- Kagan, J. C., Stein, M. P., Pypaert, M. Roy, C. R (2004). Legionella subvert the functions of Rab1 and Sec22b to create a replicative organelle. *J. Exp. Med.* 199, 1201–1211.
- Koskiniemi S, Sun S, Berg OG, Andersson DI (2012). Selection-Driven Gene Loss in Bacteria. *PLoS Genet* 8(6).
- Krinos C, High AS, Rodgers FG (1999). Role of the 25 kDa major outer membrane protein of *Legionella pneumophila* in attachment to U-937 cells and its potential as a virulence factor for chick embryos. *J Appl Microbiol* 86(2):237–244
- Kubori T, Nagai H (2016). The Type IVB secretion system: an enigmatic chimera. *Curr Opin Microbiol* 29:22–29.
- Kubori T, Shinzawa N, Kanuka H, Nagai H (2010). *Legionella* metaeffector exploits host proteasome to temporally regulate cognate effector. *PLoS Pathog* 6: e1001216. doi: 10.1371/journal.ppat.1001216 PMID: 21151961
- Kulkarni P, Cui X, Williams J, Stevens A, Kulkarni R (2006). Prediction of CsrA-regulating small RNAs in bacteria and their experimental verification in *Vibrio fischeri*. *Nucleic Acids Res* 34:3361–3369

- Kumar A, Beloglazova N, Torma C.T. et al (2016). Conditional Epistatic Interaction Maps Reveal Global Functional Rewiring of Genome Integrity Pathways in *Escherichia coli. Cell Rep.* 14(3):648-661.
- Kunert A, Hagemann M, Erdmann N. (2000). Construction of promoter probe vectors for Synechocystis sp. PCC 6803 using the light-emitting reporter systems Gfp and LuxAB**. *J. Microbiol. Methods***. 41 : 185-194
- Laguna, R. K., Creasey, E. A., Li, Z., Valtz, N. Isberg, R. R. A (2006). *Legionella pneumophila*-translocated substrate that is required for growth within macrophages and protection from host cell death. *Proc. Natl Acad. Sci. USA* 103, 18745–18750.
- Lapouge, K., M. Schubert, F. H. Allain, and D. Haas. (2008). Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behavior. *Mol. Microbiol*. 67:241–253.
- Lau H.Y, Ashbolt N.J. (2009). The role of biofilms and protozoa in *Legionella* pathogenesis: implications for drinking water. *J Appl Microbiol.* 107(2):368-78
- Leclerc, G. J., C. Tartera, and E. S. Metcalf. (1998). Environmental regulation of *Salmonella typhi* invasion-defective mutants. *Infect. Immun*. 66:682–691.
- Lee JH, Federoff H., Schoeniger LO. (1999). G207, modified herpes simplex virus type 1, kills human pancreatic cancer cells in vitro *J. Gastrointest. Surg*., 3 : 127-131
- Liles MR, Edelstein PH, Cianciotto NP (1999). The prepilin peptidase is required for protein secretion by and the virulence of the intracellular pathogen *Legionella pneumophila. Mol Microbiol* 31:959–970
- Liles, M. R., Viswanathan, V. K., and Cianciotto, N. P. (1998). Identification and temperature regulation of *Legionella pneumophila* genes involved in type IV pilus biogenesis and type II protein secretion. *Infect. Immun.* 66, 1776- 1782.
- Liu M, Haenssler E, Uehara T, Losick VP, Park JT, et al. (2012). The *Legionella pneumophila* EnhC protein interferes with immune stimulatory muramyl peptide production to evade innate immunity. *Cell Host Microbe* 12: 166-176.
- Liu Y, Luo ZQ (2007). The Legionella pneumophila effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. *Infect Immun.* 75(2):592-603
- Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2–DDCT method. *Methods* 25, 402–408

- Luo ZQ, Farrand SK (1999). Signal-dependent DNA binding and functional domains of the quorum-sensing activator TraR as identified by repressor activity. *Proc Natl Acad Sci USA* . 96:9009-9014.
- Luo, Z. Q, and Isberg R.R. (2004). Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer. *Proc. Natl. Acad. Sci. USA* 101:841–846.
- Lynch D, Fieser N, Gloggler K, Forsbach-Birk V, Marre R (2003). The response regulator LetA regulates the stationary-phase stress response in *Legionella pneumophila* and is required for efficient infection of *Acanthamoeba castellanii. FEMS Microbiol Lett* 219:241–248
- Marchal K, De Keersmaecker S, Monsieurs P, van Boxel N, Lemmens K, Thijs G, Vanderleyden J, De Moor B (2004). In silico identification and experimental validation of PmrAB targets in *Salmonella typhimurium* by regulatory motif detection. *Genome Biol* 5:R9
- Marra, A., S. J. Blander, M. A. Horwitz, and H. A. Shuman. (1992). Identification of a *Legionella pneumophila* locus required for intracellular multiplication in human macrophages. *Proc. Natl. Acad. Sci. U. S. A.* 89:9607–9611.
- Marston, B. J., J. F. Plouffe, T. M. File, Jr., B. A. Hackman, S. J. Salstrom, H. B. Lipman, M. S. Kolczak, and R. F. Breiman. (1997). Incidence of community-acquired pneumonia requiring hospitalization. Results of a population-based active surveillance study in Ohio. The Community-Based Pneumonia Incidence Study Group. *Arch. Intern. Med.* 157:1709–1718.
- Massé E, Vanderpool CK, Gottesman S (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli. J Bacteriol* 187:6962–6971
- McDade JE, Shepard CC, Fraser DW, Tsai TR, Redus MA, Dowdle WR (1977). Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. *N Engl J Med* 297:1197–1203
- McNealy, T. L., Forsbach-Birk, V., Shi, C., and Marre, R. (2005). The Hfq homolog in *Legionella pneumophila* demonstrates regulation by LetA and RpoS and inter-acts with the global regulator CsrA. *J. Bacteriol.* 187, 1527-1532

- Mendis N, McBride P, Faucher SP (2015). Short-Term and Long-Term Survival and Virulence of *Legionella pneumophila* in the Defined Freshwater Medium Fraquil. *PLoS One* 10(9).
- Mendis N, McBride P, Saoud J, Mani T, Faucher SP. (2018). The LetA/S two-component system regulates transcriptomic changes that are essential for the survival of *Legionella pneumophila* in water. *Scientific Reports*, 8: 8764.
- Merighi M, Ellermeier CD, Slauch JM, Gunn JS (2005). Resolvase-in vivo expression technology analysis of the *Salmonella enterica* serovar Typhimurium PhoP and PmrA regulons in BALB/ c mice. *J Bacteriol* 187:7407–7416
- Merrell, D. S., A. D. Tischler, S. H. Lee, and A. Camilli. (2000). Vibrio cholerae requires *rpoS* for efficient intestinal colonization. *Infect. Immun*. 68:6691–6696.
- Michod RE, Bernstein H, Nedelcu AM (2008). Adaptive value of sex in microbial pathogens. *Infection Genetics and Evolution*. 8(3):267
- Miller MB, Skorupski K, Lenz DH, Taylor RK, Bassler BL (2002). Parallel quorum sensing systems converge to regulate virulence in *Vibrio cholerae*. *Cell* 110:303–314
- Mizuno T, Chou M, Inouye M (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proc *Natl Acad Sci* 81:1966–1970.
- Molofsky A, Swanson M (2004). Differentiate to thrive: lessons from the *Legionella pneumophila* life cycle. *Mol Microbiol*.. 53:29–40
- Molofsky AB, Swanson MS (2003). *Legionella pneumophila* CsrA is a pivotal repressor of transmission traits and activator of replication. *Molecular Microbiology* 50(2):445-461.
- Morales, V. M., Backman, A. Bagdasarian, M. (1991). A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants. *Gene* 97, 39–47.
- Morris, G. K., C. M. Patton, J. C. Feeley, S. E. Johnson, G. Gorman, W. T. Martin, P. Skaliy, G. F. Mallison, B. D. Politi, and D. C. Mackel. (1979). Isolation of the Legionnaires' disease bacterium from environmental samples. *Ann. Intern. Med*. 90:664–666.
- Murata, T., A. Delprato, A. Ingmundson, D. K. Toomre, D. G. Lambright, and C. R. Roy. (2006). The *Legionella pneumophila* effector protein DrrA is a Rab1 guanine nucleotide-exchange factor. *Nat. Cell Biol.* 8:971–977.

- Murga, R., T. S. Forster, E. Brown, J. M. Pruckler, B. S. Fields, and R. M. Donlan. (2001). Role of biofilms in the survival of *Legionella pneumophila* in a model potable-water system. Microbiology 147:3121–3126
- Nagai H, Kagan JC, Zhu X, Kahn RA, Roy CR (2002). A bacterial guanine nucleotide exchange factor activates ARF on *Legionella* phagosomes. *Science* 295:679–682.
- Nakayama, S., and H. Watanabe. (1998). Identification of CpxR as a positive regulator essential for expression of the *Shigella* sonnei virF gene. *J. Bacteriol*.180:3522–3528
- Narra HP, Ochman H (2006). Of what use is sex to bacteria? *Current Biology*. 16(17):R705±R10. https:// doi.org/10.1016/j.cub.2006.08.024 PMID: 16950097
- Newton H.J, Ang D.K.Y, Driel I. R, Hartland E.L (2010). Molecular Pathogenesis of Infections Caused by *Legionella pneumophila*. *CLINICAL MICROBIOLOGY REVIEWS*. 23 (2) 274–298.
- Newton HJ, Sansom FM, Dao J, Cazalet C, Bruggemann H, Albert-Weissenberger C et al (2008). Significant role for ladC in initiation of *Legionella pneumophila* infection. *Infect Immun* 76(7):3075–3085.
- Newton HJ, Sansom FM, Dao J, McAlister AD, Sloan J, et al. (2007). Sel1 repeat protein LpnE is a *Legionella pneumophila* virulence determinant that influences vacuolar trafficking. *Infect Immun* 75: 5575-5585.
- Nomura ., Takakura , Hashida M. (1997) Cancer gene therapy by direct intratumoral injection: Gene expression and intratumoral pharmacokinetics of plasmid DNA . *Gan To Kagaku Ryoho*, 24: 483-488
- Oliva G, Sahr T, Rolando M, Knoth M, Buchrieser C. (2017). A unique cis-encoded small noncoding RNA is regulating Legionella pneumophila Hfq expression in a life cycle dependent manner. *mBio 8*:e02182-16. https:// doi.org/10.1128/mBio.02182-16.
- Palusinska-Szysz, M., Janczarek, M, Kalitynski, R., Dawidowicz, A. L, and Russa, R. (2011). *Legionella bozemanae* synthesizes phosphatidylcholine from exogenous choline. *Microbiol. Res* 166, 87-98
- Papenfort K, Vogel J (2010). Regulatory RNA in bacterial pathogens. *Cell Host & Microbe* 8:116–127.

- Phin N, Parry-Ford F, Harrison T, Stagg HR, Zhang N, Kumar K, et al (2014). Epidemiology and clinical management of Legionnaires' disease. *Lancet Infect Dis*. 14(10):1011–21.
- Piazza A, Heyer WD. (2018). Homologous Recombination and the Formation of Complex Genomic Rearrangements. *Trends Cell Biol*. 29 (2):135-149.
- Popovic D, Vucic D, Dikic I. Ubiquitination in disease pathogenesis and treatment (2014). *Nat Med* 20:1242-1253
- Prescott EM, Proudfoot NJ (2002). Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci U S A*. 99:8796–8801
- Qiu J, Luo ZQ (2017). Legionella and Coxiella effectors: strength in diversity and activity. *Nat Rev Microbiol.* 15(10):591-605.
- Qiu J, Sheedlo MJ, Yu K, Tan Y, Nakayasu ES, Das C, Liu X, Luo ZQ (2016). Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nat*. 5;533(7601):120-4.
- Raivio, T. L., and T. J. Silhavy. (1997). Transduction of envelope stress in *Escherichia coli* by the Cpx two-component system. *J. Bacteriol*. 179:7724– 7733.
- Rasis M, Segal G (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* 72:995–1010.
- Riboldi-Tunnicliffe A, König B, Jessen S, Weiss MS, Rahfeld J, et al. (2001). Crystal structure of Mip, a prolylisomerase from *Legionella pneumophila. Nat Struct Biol* 8: 779-783.
- Rodgers, F. G., P. W. Greaves, A. D. Macrae, and M. J. Lewis. (1980). Electron microscopic evidence of flagella and pili on *Legionella pneumophila. J. Clin.Pathol.* 33:1184–1188
- Romby P, Vandenesch F, Wagner EG (2006). The role of RNAs in the regulation of virulence-gene expression. *Curr Opin Microbiol* 9: 229–236.
- Romeo T (1998). Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol Microbiol* 29:1321–1330

- Romeo T, Gong M, Liu M, Brun-Zinkernagel A (1993). Identification and molecular characterization of csrA, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* 15:4744–4755.
- Rossier O, Starkenburg S, Cianciotto NP (2004). *Legionella pneumophila* type II protein secretion promotes virulence in the A/J mouse model of Legionnaires' disease pneumonia. *Infect Immun* 72:310–321
- Rothmeier E, Pfaffinger G, Hoffmann C et al (2013). Activation of Ran GTPase by a *Legionella* effector promotes microtubule polymerization, pathogen vacuole motility and infection. *PLoS Pathog* 9:e1003598. https://doi.org/10.1371/journal.ppat.1003598
- Rowbotham, T. J. (1986). Current views on the relationships between amoebae, legionellae and man. *Isr. J. Med. Sci*. 22, 678–689.
- Roy, C. R., and L. G. Tilney. (2002). The road less traveled: transport of *Legionella* to the endoplasmic reticulum. *J. Cell Biol.* 158:415–419.
- Sadosky, A. B., Wiater, L. A. & Shuman, H. A. (1993). Identification of *Legionella pneumophila* genes required for growth within and killing of human macrophages. *Infect Immun* 61, 5361–5373.
- Sahr T, Rusniok C, Ravault D.D, Sismeiro O, Coppee YJ, Buchriese C. (2009 a). Deep sequencing defines the transcriptional map of *L. pneumophila* and identifies growth phase-dependent regulated ncRNAs implicated in virulence RNA *Biology* 9:4, 503-519
- Sauvonnet N, Pugsley AP. (1996). Identification of two regions of Klebsiella oxytoca pullulanase that together are capable of promoting beta-lactamase secretion by the general secretory pathway *Mol. Microbiol*. 22 : 1-7
- Segal G, Shuman HA (1999). *Legionella pneumophila* utilizes the same genes to multiply within Acanthamoeba castellanii and human macrophages. *Infect Immun* 67(5):2117–2124
- Segal G. (2013). The *Legionella pneumophila* Two-Component Regulatory Systems that Participate in the Regulation of Icm/Dot Effectors. In: Hilbi H, editors Molecular Mechanisms in *Legionella Pathogenesis*. Heidelberg 2014
- Schousboe, A. Widmer, J. Summersgill, T. File, C. M. Heath, D. L. Paterson, and A. Chereshsky. (2002). Distribution of *Legionella* species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: an international collaborative survey. *J. Infect*. Dis. 186:127–128

- Sethi, K. K., and H. Brandis. (1983). Direct demonstration and isolation of *Legionella pneumophila* (serogroup 1) from bathroom water specimens in a hotel. *Zentralbl. Bakteriol. Mikrobiol. Hyg. B* 177:402–405.
- Sexton JA, Vogel JP (2004). Regulation of hyper competence in Legionella pneumophila. *J Bacteriol*186(12):3814-3825.

  Stone BJ, Kwaik YA (1999). Natural competence for DNA transformation by *Legionella pneumophila* and its association with expression of type IV pili. *J Bacteriol* 181(5):1395-1402.
- Sheedlo MJ, Qiu J, Tan Y, Paul LN, Luo ZQ, Das C (2015). Structural basis of substrate recognition by a bacterial deubiquitinase important for dynamics of phagosome ubiquitination. *Proc Natl Acad Sci U S A. ;*112(49):15090-5.
- Shevchuk O, Jäger J, Steinert M (2011). Virulence properties of the Legionella pneumophila cell envelope. *Front Microbiol* 2: 74.
- Sievers S, Lund A, Menendez-Gil P, Nielsen A, Storm Mollerup M, Lambert Nielsen S, Buch Larsson P, Borch-Jensen J, Johansson J, Kallipolitis BH (2015). The multicopy sRNA LhrC controls expression of the oligopeptide- binding protein OppA in Listeria monocytogenes. *RNA Biol* 12:985-97; PMID:26176322; http://dx.doi.org/10.1080/
- Sievers S, Sternkopf Lillebaek EM, Jacobsen K, Lund A, Mollerup MS, Nielsen PK, Kallipolitis BH (2014). A multicopy sRNA of Listeria monocytogenes regulates expression of the virulence adhesin LapB. *Nucleic Acids Res* 42:9383-98; PMID:25034691; http://dx.doi.org/ 10.1093/nar/gku630.
- Silhavy, T. J., Kahne, D., and Walker, S. (2010). The bacterial cell envelope *Cold Spring Harb. Perspect. Biol.* 2, a000414.
- Simons RW, Kleckner N (1983). Translational control of IS10 transposition. *Cell* 34:683–691
- Sittka, A., Lucchini, S., Papenfort K., Sharma, C. M., Rolle, K., Binnewies, T. T., Hinton, J. C., and Vogel, J. (2008). Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator Hfq. *PLoS Genet.* 4, e1000163.
- Soderberg, M. A., J. Dao, S. R. Starkenburg, and N. P. Cianciotto. (2008). Importance of type II secretion for survival of *Legionella pneumophila* in tap water and in amoebae at low temperatures. *Appl. Environ. Microbiol*. 74:5583–5588.

- Spirig T, Tiaden A, Kiefer P, Buchrieser C, Vorholt JA, Hilbi H (2008). The *Legionella* autoinducer synthase LqsA produces an alpha-hydroxyketone signaling molecule. *J Biol Chem* 283:18113–18123
- Spitalny, K. C., R. L. Vogt, L. A. Orciari, L. E. Witherell, P. Etkind, and L. F. Novick. (1984). Pontiac fever associated with a whirlpool spa. Am. *J. Epidemiol*. 120:809–817.
- Steinert M, Hentschel U, Hacker J (2002). *Legionella pneumophila*: an aquatic microbe goes astray. *FEMS Microbiol Rev* 26:149–162
- Steinert, M., and K. Heuner. (2005). Dictyostelium as host model for pathogenesis. *Cell. Microbiol*. 7:307–314.
- Stone BJ, Abu Kwaik Y (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
- Swanson MS, Hammer BK. (2000). *Legionella pneumophila* pathogesesis: a fateful journey from amoebae to macrophages. *Annu Rev Microbiol* 54: 567–613.
- Swart, A. L, Harrison, C. F, Eichinger, L, Steinert. M, Hilbi. H. (2018). *Acanthamoeba* and *Dictyostelium* as Cellular Models for *Legionella* Infection. *Frontiers in cellular and infection microbiology*, *8*, 61.
- Tanner J, Li L, Faucher SP, Brassinga AK. (2016). The CpxRA two-component system contributes to *Legionella pneumophila* virulence. *Molecular Microbiology*, 100: 1017
- Tiaden A, Spirig T, Sahr T, Walti MA, Boucke K, Buchrieser C, Hilbi H (2010). The autoinducer synthase LqsA and putative sensor kinase LqsS regulate phagocyte interactions, extracellular filaments and a genomic island of *Legionella pneumophila. Environ Microbiol* 12:1243–1259
- Tiaden A, Spirig T, Sahr T, Walti MA, Boucke K, Buchrieser C, Hilbi H (2010). The autoinducer synthase LqsA and putative sensor kinase LqsS regulate phagocyte interactions, extracellular filaments and a genomic island of *Legionella pneumophila. Environ Microbiol* 12:1243–1259
- Trigui H, Mendis N, Li L, Saad M, Faucher SP (2013). Facets of Small RNA-Mediated Regulation in *Legionella pneumophila*. In: Hilbi H, editors Molecular Mechanisms in *Legionella Pathogenesis. Heidelberg* 2014. 53-80

- Updegrove TB, Zhang A, Storz G. (2016). Hfq: the flexible RNA matchmaker. *Curr Opin Microbiol* 30:133–138. https://doi.org/10.1016/ j.mib.2016.02.003
- Vogel, J. P., Andrews, H. L., Wong, S. K. Isberg, R. R (1998). Conjugative transfer by the virulence system of *Legionella pneumophila. Science* 279, 873–876.
- Wagner C, Khan AS, Kamphausen T, Schmausser B, Unal C, et al. (2007). Collagen binding protein Mip enables Legionella pneumophila to transmigrate through a barrier of NCI-H292 lung epithelial cells and extracellular matrix. *Cell Microbiol* 9: 450-462.
- Washington C, Winn JR. Medical Microbiology. 4th edn. (1996). University of Texas Medical Branch at Galveston, Galveston, Texas
- Wassarman KM (2007). 6S RNA: A regulator of transcription. *Mol Microbiol* 65: 1425–1431.
- Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 136:615–628
- Weissenmayer B.A, Prendergast JG, Lohan A.J, Loftus B.J. (2011). Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs. *PLoS One* 6(3): e17570.
- Wood, R. E, Newton, P, Latomanski, E. A, Newton, H. J. (2015). Dot/Icm Effector Translocation by *Legionella longbeachae* Creates a Replicative Vacuole Similar to That *of Legionella pneumophila* despite Translocation of Distinct Effector Repertoires. *Infection and immunity*, *83*(10), 4081–4092.
- Zhan X. Y, Hu C.H, Zhu Q. Y. Legionella Pathogenesis and Virulence Factors (2015). *Annals of Clinical and Laboratory Research.* 3 (2): 15
- Zhou Y, Zhu Y (2015). Diversity of bacterial manipulation of the host ubiquitin pathways. *Cell Microbiol* 17(1):26–34.
- Zusman T, Aloni G, Halperin E, Kotzer H, Degtyar E, Feldman M, Segal G (2007). The response regulator PmrA is a major regulator of the *icm/dot* type IV secretion system in *Legionella pneumophila* and *Coxiella burnetii*. *Mol Microbiol* 63:1508–1523.