The Dissertation Committee for Casey Lee Cody Schroeder Certifies that this is the approved version of the following dissertation:

Rickettsia prowazekii and its 'Junk DNA': The Identification and Characterization of Small RNAs

Committee:

Sanjeev K. Sahni, PhD, Supervisor

Yuriy Fofanov, PhD

Michael Minnick, PhD

Gustavo Valbuena, MD, PhD

David Walker, MD

Dean, Graduate School

Rickettsia prowazekii and its 'Junk DNA': The Identification and Characterization of Small RNAs

by

Casey Lee Cody Schroeder, MS, MLS(ASCP), SM, SLS

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas Medical Branch

October, 2016

Dedication

I dedicate this dissertation to my daughter, who has the fortunate pleasure of possessing half of her mother's genes. Sorry about the other half.

Explore. Dream. Discover.

For I learned far more in the woods than I did inside a building.

Nature is calling. Answer it.

Acknowledgments

I would like to acknowledge those that were instrumental:

-*To the Flying Spaghetti Monster:* For your noodly appendage has touched and guided my results. Without the logical conjecture based on overwhelming observable evidence, the conclusions in this dissertation would not have been possible.

-*The Universe*: For had you not expanded 13.8 billion years ago and created the Earth 4.5 billion years ago, none of this research would have been physically possible. The continued spewing of chemically-rich star guts atomically binds this research to the universe and biologically to the Earth. For that, I am eternally grateful.

-*Dissertation Committee*: My sympathies for forcing you to sit through, what probably felt like never-ending, biannual dissertation committee meetings with me. If it is any solace, they have ended……. hopefully. Specifically, Dr. Fofanov, you're fired. Dr. Minnick, enjoy fly fishing with the grizzly bears. Dr. Valbuena, sweet job in California. Dr. Walker, live long and prosper. Dr. Sahni, I know, it's about time.

-*Evolution*: For you are the greatest show on Earth!

-*Space:* For Douglas Adams said it best, you are big. Really big. Vastly, hugely, mind-bogglingly big. Thanks for that.

-*Homo neanderthalensis*: Hell of a thing 40,000 years ago. On behalf of all *Homo sapiens*, sorry it did not work out. If it is any consolation, your demise allowed *Homo sapiens* to discover science and the scientific method. Without that, this dissertation would have never been accomplished.

-*Rickettsia prowazekii*: It is truly unfortunate that I was required to sacrifice some of your living descendants. I am greatly saddened by their loss. Nevertheless, their contributions to this work and to rickettsiology in general will not be forgotten.

-Don't Panic! So long, and thanks for all the fish!

-Oh, and SCIENCE RULES!

Rickettsia prowazekii and its 'Junk DNA': The Identification and Characterization of Small RNAs

Publication No._____

Casey L. C. Schroeder, MS, PhD, MLS(ASCP), SM, SLS

The University of Texas Medical Branch, 2016

Supervisor: Sanjeev K. Sahni, PhD

As the etiologic agents of Rocky Mountain spotted fever and epidemic typhus, *Rickettsia* are obligate intracellular Gram-negative pathogenic bacteria. Genomic sequencing determined that ~24% of the *Rickettsia prowazekii* genome was noncoding DNA due to genomic reduction. Regions of bacterial noncoding DNA were found to encode important post-transcriptional regulators of bacterial virulence and growth called bacterial small RNAs (sRNAs). These sRNAs are classified as either *trans*-acting, whose biogenesis is predominantly linked to intergenic regions, or *cis*-acting, encoded on the antisense strand of an open reading frame (ORF). Despite being present in a majority of other bacteria, the existence and function of sRNAs are unknown in *Rickettsia*. In order to explore the possibility of rickettsial sRNAs, a combination of bioinformatics and *in vitro* techniques was employed. The first objective identified over 1,700 sRNAs using the SIPHT analysis tool to analyze 16 different strains representing 13 rickettsial species. Strong σ70 promoters and Rho-independent terminators were detected in their respective positions for all candidate sRNAs predicted in *R. prowazekii, R. typhi*, *R. conorii,* and *R. rickettsii*. Next generation sequencing (NGS) performed at 3h and 24h on *R. prowazekii*-

infected human microvascular endothelial cells (HMECs) validated the expression of 26 sRNA candidates. Six selected candidates were further confirmed using RT-PCR. The second objective analyzed the NGS data and found the expression of an additional 35 *trans*- and 23 *cis*-acting novel sRNAs in addition to well-conserved kingdom-wide sRNAs (6.5S, 4.5S, RNaseP_bact_a, α-tmRNA). Three novel sRNAs and the conserved sRNAs were confirmed using either Northern blot analysis or RT-PCR. Transcriptional start sites were determined using RLM-RACE and secondary structure predicted for five novel sRNAs and 6S RNA. The third objective examined comparison of sRNA expression between HMECs and AAE2 (arthropod host cells), and found that using qRT-PCR, four sRNAs that showed significantly higher levels of expression at 24h. Also, a closer examination of *Rp*_sR60 found an interaction with H375_0420, in which both sRNA and target were found to be significantly decreased at 3h and 24h in HMECs. During AAE2 infection, both of these targets were found to be consistently expressed at 0.5h, 3h, and 24h. Together, these results establish the presence and expression of sRNAs in *R. prowazekii* during host cell infection and suggest potential functional roles for these important post-transcriptional regulators in rickettsial biology and pathogenesis.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

AAE2	*Amblyomma americanum* tick cell line
BLAST	Basic Local Alignment Search Tool
BPROM	Bacterial Promoter Program
BSL	Biosafety Level
CTP	Carboxyl Terminal Protease
DNA	Deoxyribonucleic Acid
HGT	Horizontal Gene Transfer
HMEC	Human Microvascular Endothelial Cells
IGR	Intergenic Regions
MEV	Mean Expression Value
NGS	Next Generation Sequencing
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
qRT-PCR	Real-time, Reverse-Transcriptase Polymerase Chain Reaction
RLM-RACE	RNA Ligase Mediated Rapid Amplification of cDNA Ends
RNA	Ribonucleic Acid
RNA-seq	Next Generation RNA Sequencing
*Rp*_sR	*Rickettsia prowazekii* small RNAs
RT-PCR	Reverse-Transcriptase Polymerase Chain Reaction
SIPHT	sRNA Identification Protocol using High throughput Technology
sRNA	Small, Noncoding RNAs
T3SS	Type 3 Secretion System
T4SS	Type 4 Secretion System

Chapter 1 Introduction to *Rickettsia* and Bacterial Small RNAs[1]

INTRODUCTION

The Genus *Rickettsia* includes Gram-negative bacilli that belong to the class *α-proteobacteria*. Relatively small AT-rich genomes, fastidious growth requirements, natural transmission through arthropod vectors, and a tropism for endothelial cells in mammalian hosts define these obligate intracellular pathogens (Figure 1) (Blanc *et al.*, 2007; Darby *et al.*, 2007; Fuxelius *et al.*, 2007). Although rickettsiae are traditionally divided into antigenically distinct groups, namely spotted fever and typhus, recent phylogenetic evidence now categorizes the genus into four main groups: ancestral, typhus, transitional, and spotted fever (Gillespie *et al.*, 2008). The ancestral group consists of *Rickettsia bellii* and *R. canadensis*. The typhus group, responsible for epidemic typhus and endemic typhus, includes only two species: *R. prowazekii* and *R. typhi*. The most diverse group in terms of vectors and disease is the transitional group, comprised of three species: *R. australis*, *R. akari*, and *R. felis*. Nevertheless, the spotted fever group includes the largest species diversification with at least 15 species, including *R. rickettsii* and *R. conorii* (Table 1). There are several other rickettsial species that have yet to be fully characterized, such as *R. asiatica*, *R. hoogstraalii*, and *R. argasii* (Fujita *et al.*, 2006; Lafri *et al.*, 2015).

[1] Parts of this introduction may appear in Schroeder C. L., I. H. Chowdhury, H. P. Narra, J. Patel, A. Sahni, and S. K. Sahni. Human rickettsioses: host response and molecular pathogenesis. Rickettsiales: Epidemiology, Molecular Biology and Vaccine Development, 2016, Springer, New York [Invited book chapter – Submitted]. Springer allows a dissertation license that permits an advanced degree candidate to republish the requested material in his/her dissertation as per their instructions_(https://s100.copyright.com/help/rightslinkhelppages/Frequently_Asked_Questions_Springer.htm)

RICKETTSIOSES

Rickettsial diseases, such as Rocky Mountain spotted fever (RMSF) and epidemic typhus, have had a significant impact on nearly all facets of society and historical events. RMSF was first described in 1896 by Major Marshall Wood, U.S. Army Medical Corps, in Boise, Idaho, U.S.A. (Thorner *et al.*, 1998). At that time, RMSF was colloquially referred to as "blue disease" or "black measles" (Woodward, 1973). Later in 1906, Howard Ricketts demonstrated tick transmission of bacilli and pathology resembling RMSF in guinea pigs (Bechah *et al.*, 2008b). The disease, broadly distributed in the United States despite its name, has a geographic range as far north as Canada and as far south as Argentina. The greatest incidence of spotted fever rickettsioses in the United States occurs in a zone that stretches north-south between northern Missouri and southern Arkansas and east-west between mid-Oklahoma and the east coast (Drexler *et al.*, 2016). RMSF and other spotted fever rickettsioses are notifiable diseases in every state of the United States except Hawaii and Alaska (Drexler *et al.*, 2016). If left untreated, RMSF may have a case fatality rate between 20-25% (Smadel, 1959; Childs and Paddock, 2002). Even in the post-antibiotic era, the case fatality rate hovers around 1-2% (Openshaw *et al.*, 2010). Delayed diagnosis worsens the clinical course of RMSF as it increases the chances of complications and fatal outcome. The key to successful treatment is early administration of appropriate antibiotics (Thorner *et al.*, 1998; Drexler *et al.*, 2016). To date, 2012 has had the highest reported incidence of spotted fever rickettsioses in the United States with approximately 4,500 cases (Drexler *et al.*, 2016).

As RMSF is considered to be the most severe spotted fever rickettsiosis, the same holds true for epidemic typhus from the typhus group. Historically, epidemic typhus was

dubbed the "scourge of the armies", as it was a common infectious disease during wars, famine, and within impoverished areas (Raoult *et al.*, 2006; Zhang *et al.*, 2006; Badiaga and Brouqui, 2012). Caused by *R. prowazekii,* epidemic typhus played a significant role in decimating the strength of Napoleon's Grand Army of 1812 (Raoult *et al.*, 2004). During World War II, outbreaks of epidemic typhus were common in military camps, Nazi concentration camps, and battle torn cities. Antibodies for *R. prowazekii* were found in two homeless individuals in Houston, Texas, U.S.A in 2008. However, these individuals were not exhibiting signs or symptoms of epidemic typhus at the time of serological testing (Reeves *et al.*, 2008). More recently, outbreaks have occurred in Burundi and Russia (Badiaga and Brouqui, 2012). While there is a controversy on whether epidemic typhus originated in the "Americas" or in Europe, today the disease is found worldwide in conjunction with *Pediculus humanus corporis* (human body louse) (Bechah *et al.*, 2008a). In the pre-antibiotic era, the mortality rate for epidemic typhus reached as high as 60%. However, after the development of antibiotics, the mortality rate dropped to as low as 4%. Nevertheless, as the disease is prevalent in conditions of famine, war, population displacement, and malnutrition, the mortality rate can still reach as high as 50% (Bechah *et al.*, 2008a). Interestingly, only *R. prowazekii* has the ability to sustain subclinical latent infections in convalescent persons, who can later develop recrudescent typhus, also known as Brill-Zinsser disease. This form of epidemic typhus is generally a milder disease (Bechah *et al.*, 2008a; Sahni *et al.*, 2013). Currently, there are no licensed vaccines for *R. prowazekii* or other rickettsial diseases in the United States or elsewhere.

RICKETTSIAL TRANSMISSION AND ARTHROPOD VECTORS

Association with arthropods is responsible for natural transmission of human rickettsioses through infected vectors. In the case of *Rickettsia* species, these vectors include ticks, cat fleas, rat fleas, mites, and human body lice. The spotted fever group *Rickettsia* and *R. australis* are transmitted through ticks (Figure 2), while the remaining transitional group *Rickettsia* and typhus group *Rickettsia* are primarily transmitted through non-tick vectors. Humans are considered to be "dead-end hosts", as they are not essential to the rickettsial lifecycle, with the notable exception of *R. prowazekii*. In addition, the geographic ranges of rickettsial diseases are limited only by their natural vector's range (Azad and Beard, 1998).

Rickettsia-free ticks acquire spotted fever group *Rickettsia* when the tick takes a sufficient blood meal from a rickettsemic host (Socolovschi *et al.*, 2009). Co-feeding is another situation in which ticks acquire *Rickettsia*. This can occur when an infected tick feeds on a host in close vicinity to *Rickettsia*-free ticks. Although the host may not yet be rickettsemic, the *Rickettsia*-free ticks may acquire the pathogen from tissue being fed upon by the infected tick (Philip, 1959). *Rickettsia* have been shown to be passaged by transovarial (to offspring ticks) and transstadial (from one arthropod growth stage to the next) transmission, which ensures their natural maintenance without the need of an intermediate host (Sahni *et al.*, 2013).

The method of transmission is different for *Rickettsia* species transmitted through louse and flea vectors. *R. prowazekii* (epidemic typhus) is transmitted through *Pediculus humanus corporis* (human body louse; Figure 3), while *R. typhi* (endemic or murine typhus) is transmitted through *Xenopsylla cheopis* (Oriental rat flea). Although, the

human body louse is the known primary vector of *R. prowazekii*, there have been reports of its occurrence in *Hylomma* ticks in Ehiopia and *Amblyomma imitator* ticks from Mexico, and the ectoparasites of *Glaucomys volans volans* (eastern flying squirrel) (Philip *et al.*, 1966; Bozeman *et al.*, 1975; Medina-Sanchez *et al.*, 2005). Despite the vector, these rickettsiae are not transmitted directly during ingestion of the blood meal from the host, unlike tick-borne rickettsiae. Instead, as the arthropod vector defecates during the course of taking its blood meal, the host scratches the bite wound and accidently introduces *Rickettsia*-contaminated feces into the wound, leading to infection (Azad and Beard, 1998). While transovarial passage for flea-borne *Rickettsia* (*R. typhi* and *R. felis*) has been shown, *Rickettsia*-infected lice die within 2-weeks (Farhang-Azad *et al.*, 1985; Azad *et al.*, 1992). Therefore, rapid transmission is essential for the survival of *R. prowazekii*, which lends a possible explanation for recrudescent typhus; i.e., the patients with Brill-Zinsser disease may serve as potential host reservoirs for epidemic typhus.

GENOMICS

In 1998, the chromosome of *R. prowazekii* strain Madrid E became the first rickettsial genome to be fully sequenced and published (Andersson *et al.*, 1998). An examination of the genome revealed a pattern consistent with reductive evolution as *Rickettsia* continue to purge unnecessary and redundant genes to adapt to the obligately intracellular lifestyle. Many of the genes required by free-living bacteria are absent in *Rickettsia,* such as the biosynthetic pathways. Further, nearly 150 nucleus-encoded mitochondrial proteins are homologous to proteins found in *R. prowazekii,* with

significant similarity in transport mechanisms and ATP production associated with the TCA cycle and respiratory chain complexes. Once encoded in the mitochondrial ancestor genome, nucleus-encoded mitochondrial proteins help to control mitochondrial functions, but are now encoded in the nuclear DNA.

Data suggest that the origin of aerobic respiration may have evolved from an ancestor of *Rickettsia*, which possibly diverged approximately 1.5 to 2 billion years ago (Andersson *et al.*, 1998), followed by transition of a free-living *Rickettsiales* ancestor to the obligately intracellular lifestyle between 775-525 million years ago (Merhej and Raoult, 2011). *Rickettsiales*' association with arthropods did not occur until the Cambrian explosion, in which arthropods first evolved approximately 542 to 500 million years ago. The genus known today as *Rickettsia* did not appear until the evolution of the rickettsial hydra group 150 million years ago. A rapid radiation of rickettsial groups occurred 50 million years ago bringing about the ancestors to the rickettsial groups known today (Weinert *et al.*, 2009b).

The genus *Rickettsia* has a genome size ranging from 1.11 Mb to 2.1 Mb (Table 1). The genomes of spotted fever group *Rickettsia* range from 1.27 to 2.1 Mb, while typhus group is approximately 1.11 Mb. Ancestral group *Rickettsia* and transitional group *Rickettsia* genomes range from 1.15 to 1.53 Mb and 1.23 to 1.59 Mb, respectively (Merhej *et al.*, 2014). Surprisingly, rickettsial genomes are rather AT-rich with GC contents ranging from 28.9% in typhus group to 33.3% in the spotted fever group. Typhus group rickettsiae have the least number of annotated open reading frames (ORFs) and proteins encoded in their genomes, 865 to 999 and 829 to 950, respectively. In the course of switching from a free-living to an obligately intracellular bacterium, data

suggest that *Rickettsia* have lost about 2,135 genes due to reductive evolution (Georgiades *et al.*, 2011; Georgiades and Raoult, 2011; Merhej and Raoult, 2011). Interestingly, *Rickettsia* have a high percentage of non-coding DNA when compared to other bacteria. *Rickettisa prowazekii* and *R. conorii* contain 24% and 19% non-coding DNA, respectively, yet *Chlamydia trachomatis*, another obligate intracellular bacterium, contains only 10% non-coding DNA (Andersson *et al.*, 1998; Holste *et al.*, 2000; Rogozin *et al.*, 2002). *Carsonella ruddii*, an endosymbiont of psyllids, encodes for the smallest known genome with 182 ORFs and 97.3% coding density (Nakabachi *et al.*, 2006).

Due to the obligately intracellular lifestyle of rickettsiae, mobile genetic elements were long considered to be uncommon. Today, at least 11 *Rickettsia* species have been found to contain plasmids with diversity among the number carried by different species/strains (El Karkouri *et al.*, 2016). These 11 species, representing the transitional group (*R. australis* and *R. felis*) and the spotted fever group (*R. africae, R. amblyommii, R. helvetica, R. massiliae, R. monacensis, R. peacockii, R. raoultii,* and *R. rhipicephali*) harbor at least 20 known rickettsial plasmids (Table 1). It is important to consider, however, that nine spotted fever group, one transitional group, and two typhus group *Rickettsia* species have no detectable plasmid. These species include *R. akari, R. conorii, R. japonica, R. montanensis, R. parkeri, R. philipii, R. rickettsii, R. sibirica, R. slovaca, R. prowazekii, R. typhi,* and *R. canadensis* (El Karkouri *et al*., 2016). The first rickettsial plasmid, pRF, was discovered after the sequencing of the *R. felis* genome. Although two plasmids, pRF and pRFδ, were initially described, pRFδ is mostly likely an artifact of the genome assembly for a number of reasons that include the absence of required

maintenance genes and failure to amplify in other strains (Gillespie *et al.*, 2007). However, the 62.8 kb pRF plasmid contains 68 ORFs encoded within its sequence. Upon sequencing the REIS genome, four novel plasmids (pREIS1 through pREIS4) were identified (Gillespie *et al.*, 2012). In addition to the plasmids, the REIS genome was found to be overrun with more than 650 transposons, mobile genetic elements, and amplified genetic elements. This suggests that rickettsial genes associated with the intracellular lifestyle were acquired via gene transfer (Gillespie *et al.*, 2012). Interestingly, analysis of plasmid gene content, exhibiting 16% to 59% gene degradation, suggests that rickettsial plasmids are also undergoing reductive evolution similar to that seen on chromosomes (El Karkouri *et al.*, 2016).

Rickettsial plasmids and mobile genetic elements highlight the importance of horizontal gene transfer (HGT). Although first described in the 1940s, HGT is a critical evolutionary mechanism that ensures continued adaptation to new environments by conjugation, transformation, or transduction (Sahni *et al.*, 2013; Soucy *et al.*, 2015). Previous research into rickettsial evolution placed only scant emphasis on rickettsial HGT due to its intracellular lifestyle. Nevertheless, a focus on HGT and its role in rickettsial evolution after genome sequencing has continuously demonstrated mobile genetic elements (Ogata *et al.*, 2006; Gillespie *et al.*, 2007; Weinert *et al.*, 2009a; Merhej *et al.*, 2011; Merhej and Raoult, 2011; Gillespie *et al.*, 2012; El Karkouri *et al.*, 2016). To determine the origin of rickettsial plasmids, an examination of 260 plasmid genes was conducted. Vertical transmission of the last common plasmid ancestor, labeled pRICO, was partially responsible for the current set of rickettsial plasmids. The plasmids, despite undergoing gene loss resulting in reductive evolution, were successful in gaining new

genes via HGT and by numerous duplication events. Although plasmids gained genes from HGT, the degradation of genes probably resulted in permanent plasmid loss noted in several rickettsial species (El Karkouri *et al.*, 2016). Evidence of such HGT has been demonstrated by the observation of pili between two rickettsial bacteria, *R. felis* and *R. bellii*. For *R. felis,* two pili were observed by transmission electron microscopy on the surface of the bacteria. These include pili that are conjugative in nature and establish direct contact with other bacteria and short hair-like pili that are most likely involved in bacterial attachment to host cells (Ogata *et al.*, 2005). A *tra* gene cluster conserved in the *R. bellii* genome sequence supports its ability to form pili. *Tra* genes are composed of an F-like and a Ti-like region (Ogata *et al.*, 2006). The F-like region encodes 12 conjugative transfer proteins, which correspond to the Type IV Secretion System (T4SS) core proteins and auxiliary proteins (Lawley *et al.*, 2003). The Ti-like region encodes the TraA ORF, responsible for the activity of nickase and helicase in initiating DNA transfer, and TraD, which has an unknown function.

BACTERIAL SMALL RNAS[2,3]

Small noncoding RNAs (sRNAs) were first identified and described in the 1960s, but remained largely ignored until recently when they were recognized as important post-transcriptional regulators in both eukaryotic and prokaryotic organisms (Wassarman, 2007). Bacteria, especially intracellular species, are required to quickly adapt to ever changing and often hostile environments. Small RNAs allow for a rapid response and fine-tuning to environmental changes such as decreasing pH, increasing reactive oxygen species, poor nutrient contents, temperature, and any immune responses (Oliva *et al.*,

2015). These sRNAs have been found in a number of pathogenic bacteria belonging to family Enterobacteriaceae, *Pseudomonas aeruginosa, Listeria monocytogenes, Streptococcus pyogenes, Clostridium perfringens,* and *Staphylococcus aureus* (Papenfort and Vanderpool, 2015). For example, *Helicobacter pylori,* the causative agent of chronic active, chronic persistent, and atrophic gastritis in adults and children, and implicated in a majority of duodenal and gastric ulcers, carries a repertoire of at least one anti-sense transcriptional start site on approximately 46% of its open-reading frames, 28% of tRNAs, and the 5' leader sequences for both 16S rRNA and 23S rRNA (Sharma *et al.*, 2010).

Barring a few exceptions, sRNAs are transcripts typically 50 to 500 nucleotides in length and do not code for proteins (Xiao *et al.*, 2009; Gottesman and Storz, 2011; Stazic and Voss, 2016). Unlike riboswitches, sRNAs are independent transcripts under the control of sigma promoters and Rho-independent terminators (Otaka *et al.*, 2011). Despite being longer, these sRNAs are considered to be analogous to eukaryotic sRNAs in the context of certain functional implications. Post-transcriptional sRNA-mediated regulatory mechanisms are broadly categorized into sRNA-protein and sRNA-mRNA interactions (Waters and Storz, 2009; Wagner and Romby, 2015). Interactions with the protein-coding transcripts are further categorized into two groups, namely, trans-acting and cis-acting (Liu and Camilli, 2010; Gottesman and Storz, 2011).

By definition, *cis*-acting sRNAs are encoded on the anti-sense strand and generally display perfect nucleotide complementarity with the target sequence in the open reading frame. Due to this condition, the sRNA interacts with the mRNA transcript autonomously (Caldelari *et al.*, 2013). Generally, *cis*-acting sRNAs are located within

the untranslated regions (UTRs) of an ORF. Upon RNA duplex formation, the sRNA can alter the secondary structure of the mRNA transcript, thus preventing ribosome entry or termination. Also, the RNA duplex may prevent the ribosome from binding to the ribosome binding site (Oliva *et al.*, 2015). Either scenario would decrease translation of the transcript and ultimately the expression. Conversely, a rearrangement in secondary structure may increase stability of the mRNA and/or free the ribosome binding site. These actions would eventually increase the expression of the protein. *Cis*-acting sRNAs are found in both Gram-positive and Gram-negative organisms. The degree to which sRNA density is found within these organisms varies greatly (Raghavan *et al.*, 2012). Many *cis*-acting sRNAs lack a functional characterization making them a fascinating interest in bacterial genomics (Oliva *et al.*, 2015).

Trans-acting sRNAs, on the other hand, are encoded within the intergenic regions, act on targets elsewhere in the genome, and possess short segments of partial nucleotide complementary to their target genes (Waters and Storz, 2009; Wagner and Romby, 2015; Stazic and Voss, 2016). Accordingly, they require a known RNA chaperone, namely Hfq, encoded by nearly 50% of all bacterial species to facilitate their binding interactions with an mRNA transcript (Waters and Storz, 2009; Papenfort and Vanderpool, 2015). However, this also means that 50% of bacterial species are not known to encode an Hfq protein; instead, they presumably use another RNA chaperone. In organisms such as *Listeria monocytogenes*, most *trans*-acting small RNAs function independently of Hfq chaperone activity (Nielsen *et al.*, 2011). Alternative proteins that function as RNA chaperones have been proposed in *Helicobacter pylori* and *Mycobacterium tuberculosis.* For *H. pylori,* the protein HP1334 stably binds to HPnc6910 sRNA, whereas *GroEL1* in

M. tuberculosis has been shown to bind both dsDNA and ssDNA, suggesting a potential chaperone function (Basu *et al.*, 2009; Rieder *et al.*, 2012). The genus *Rickettsia* is not known to encode an Hfq protein, nor has an alternative chaperone been described.

Historically, the majority of intergenic regions has traditionally been considered 'junk DNA' or defunct genes resulting from reductive evolution. The existence of sRNAs and their role in transcriptome regulation in *Rickettsia* spp. and other obligately intracellular bacteria remain largely unknown. The central hypothesis of this work is that since genomic reduction, the main driving force shaping rickettsial genomes, results in maintaining only a minimal set of genes required for survival and the obligately intracellular life style. The *Rickettsia* genome encodes small RNAs required for regulation of its transcriptome, and differentially expresses these sRNAs leading to altered virulence and adaptation depending on the host niche.

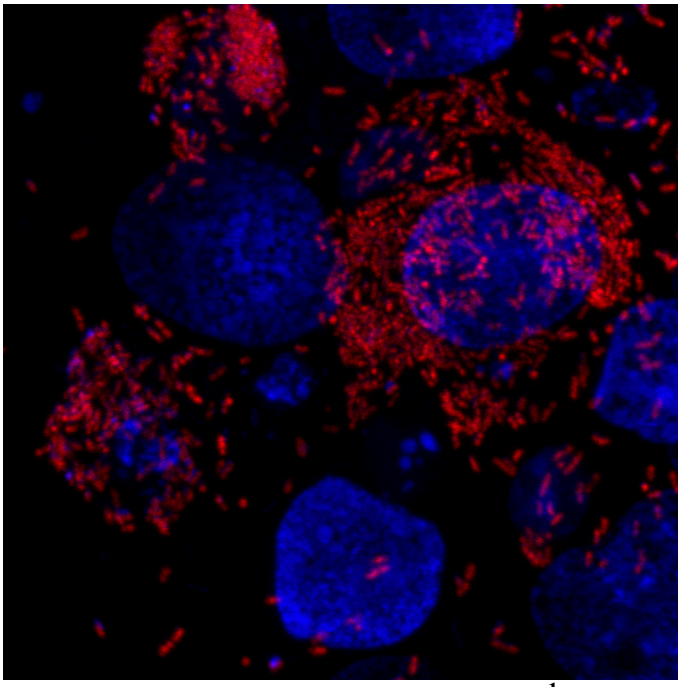


Figure 1. *Rickettsia conorii* in host cells.[1]

Confocal laser scanning microscopy image of Vero cells infected with *Rickettsia conorii* expressing kusabira orange. DAPI (4',6'-diamindino-2-phylindole) was used for fluorescent staining (blue) of host cell nuclei.

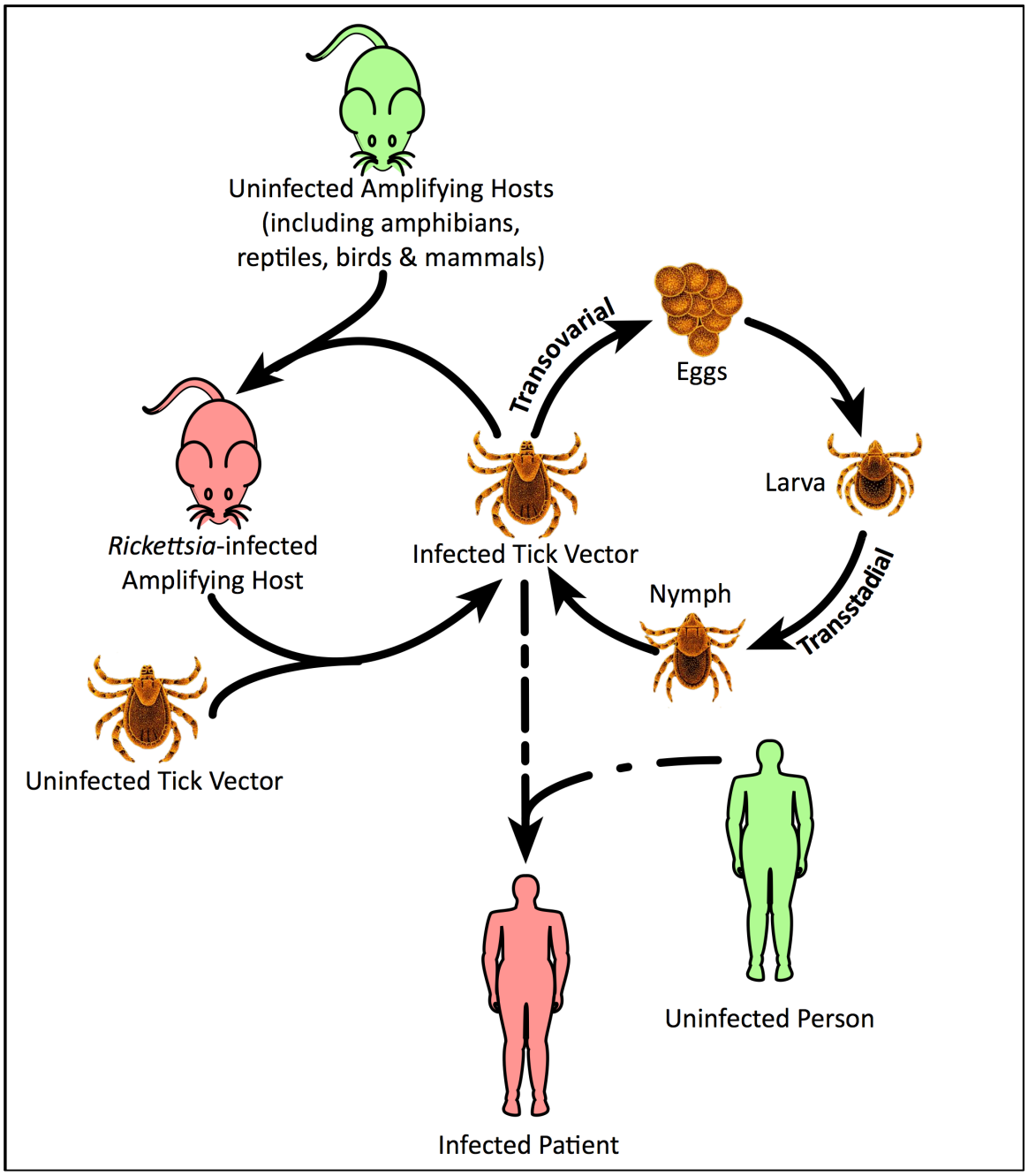


Figure 2. Transmission cycle for the natural maintenance of spotted fever group *Rickettsia*.[1]

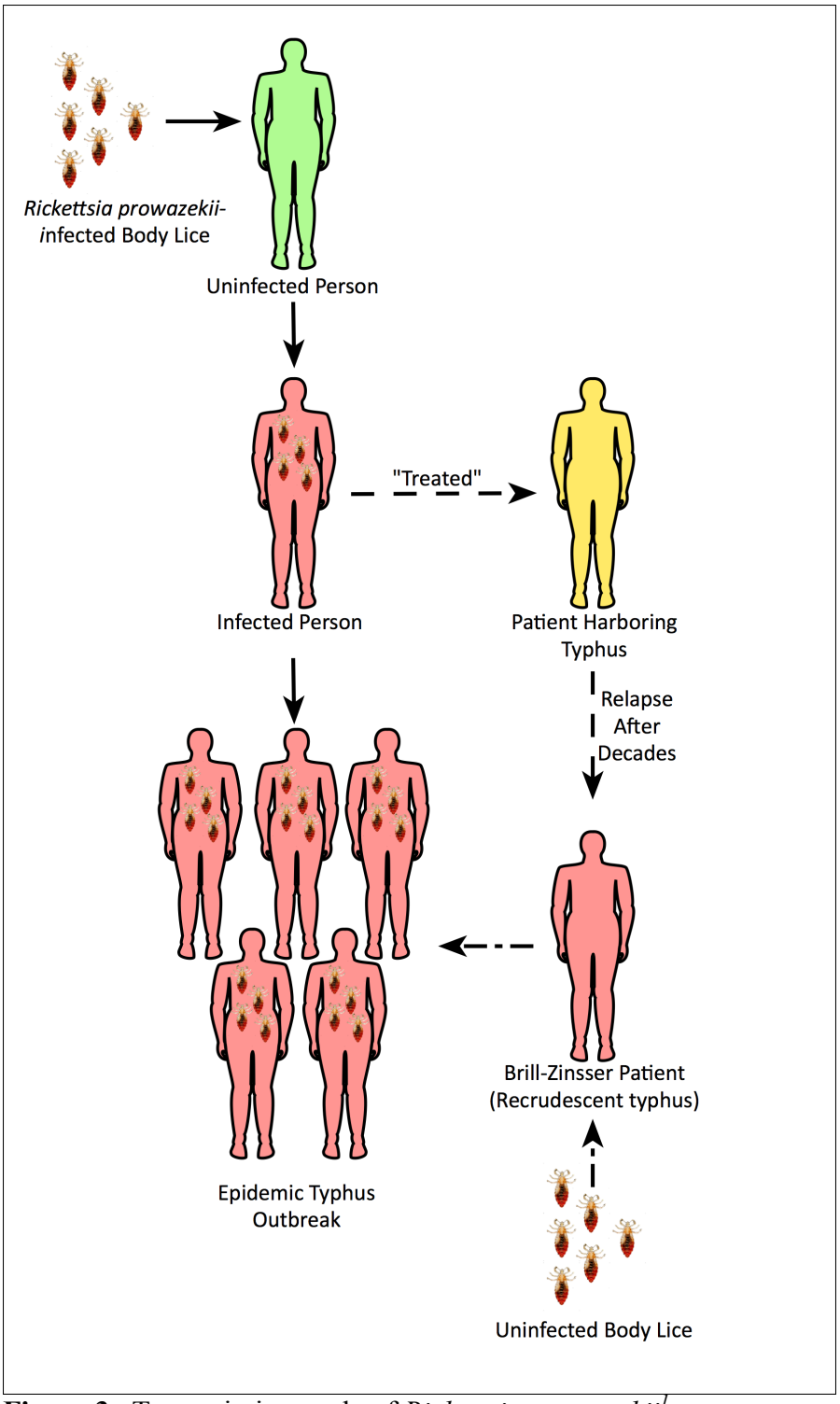


Figure 3. Transmission cycle of *Rickettsia prowazekii*[1]

Table 1. Rickettsial genomes and associated plasmids.

Rickettsial Group	Species	Strain	Genome Size (bp)	GC%	CDS	Plasmid
Ancestral Group						
	R. bellii					
		RML 369-C	1,552,076	31.6	1429	-
		OSU 85-389	1,528,980	31.6	1476	-
	R. canadensis	McKiel	1,159,772	31.1	1093	-
Typhus Group						
	R. prowazekii					
		Madrid E	1,111,523	29	835	-
		Rp22	1,111,612	29	952	-
		Breinl	1,109,301	29	920	-
	R. typhi	Wilmington	1,111,496	28.9	838	-
Transitional Group						
	R. akari	Hartford	1,231,060	32.3	1259	-
	R. australis					
		Cutlack	1,323,280	32.3	1261	pRau
		Phillips	1,320,570	32.2	1715	-
	R. felis	URRWXCal2	1,485,147	32.5	1512	pRfe pRfel1
Spotted Fever Group						
	Candensis R. amblyommii		1,448,020	32.4	1821	pRam18 pRam23 pRam32
	R. africae	ESF-5	1,278,540	32.4	1041	pRaf
	R. conorii	Malish 7	1,268,755	32.4	1374	-
	R. helvetica	C9P9	1,369,827	32.2	1739	pRhe
	R. honei	RB	1,268,760	32.4	1614	-
	R. japonica	YH	1,279,890	32.4	971	-
	R. massiliae					
		MTU5	1,376,180	32.5	980	pRma
		AZT80	1,278,720	32.6	1207	pRmaB
	R. monacensis	IrR/Munich	1,353,450	32.4	1460	IrR/Munich
	R. parkeri	Portsmouth	1,300,386	32.4	1318	-
	R. peacockii	Rustic	1,288,492	32.6	947	pRpe
	R. raoultii					pRra1 pRra2 pRra3
	R. rhipicephali					
		3-7-female6-CWPP	1,305,470	32.4	1266	pRrh
		HJ#5	1,448,630	32.3	1255	-
		Ect	1,266,920	32.6	1563	-
	R. rickettsii					
		Sheila Smith	1,257,710	32.5	1345	-
		Iowa	1,268,201	32.4	1384	-
		R	1,257,005	32.5	1334	-
		Brazil	1,255,681	32.5	1332	-
	R. sibirica					
		mongolitimonae	1,252,340	32.4	1616	-
		sibirica BJ-90	1,254,730	32.5	1588	-
	R. slovaca					
		13-B	1,275,090	32.5	1112	-
		D-CWPP	1,275,720	32.5	1347	-
	REIS		2,096,878	33	2117	pReis1 pReis2 pReis3 pReis4

Chapter 2 Methodology

RICKETTSIA SPECIES AND STRAINS

For this study, available genome sequences of 16 rickettsial strains, encompassing 13 species belonging to the genus *Rickettsia* were used. Included species represent all four, namely ancestral, typhus, transitional, and spotted fever, rickettsial groups (Table 2). Further, four known rickettsial plasmids were also analyzed (Table 2).

PREDICTION OF sRNAs

I used the web-based program SIPHT available from the University of Wisconsin at Madison (http://newbio.cs.wisc.edu/sRNA/index.php) for predicting sRNAs in rickettsial genomes (Livny *et al.*, 2008). This program predicts sRNAs within the intergenic regions of bacterial genomes by searching for Rho-independent terminators downstream of conserved sequences, followed by an analysis of conservation with other species, potential transcription factor binding sites, the spacing between flanking genes, and homology with known sRNAs (Livny *et al.*, 2008). The specific parameters used for all searches were as follows. The maximum expected value for BLAST (BLAST E) was changed from the default setting of 5e-3 to a more stringent value of 5e-15 in order to eliminate the possibility of false positives. The minimum score for BLAST (BLAST S) and minimum percent identity (BLAST % identity) were set at the default 0. The maximum BLAST high-scoring segment pairs (HSP) length was set at 1000. The maximum Rho-independent terminator criteria were 86, -10, and -6 for TransTerm, FindTerm, and RNAMotif, respectively. The minimum predicted locus length was 30,

while the maximum predicted locus length was 550. These scores take into account the generally defined 50 to 500-nucleotide length of bacterial sRNAs (Gottesman and Storz, 2011). The minimum distance, by default, for both the locus start site to ORF start site and for the locus start site to ORF end site was -65. However, the minimum distance set by default from the locus end site to ORF start site was -20. The minimum distance from the locus end site to ORF stop site was left at 35. Lastly, the minimum distance from the transcription factor-binding site to the ORF start site was 0. This value allows for all possible candidate sRNAs to be included in the final report. To ensure consistency, all of these parameters were set as the default analytical criteria for the program.

PROMOTER PREDICTION

For bacterial promoter predictions, the web-based software BPROM was used. This program searches for bacterial σ^{70}-family promoter -10 box and -35 box, transcription start site, and other transcription factor binding sites in a given genomic sequence with a reported accuracy and specificity of 80% (Solovyey and Salamov, 2010). Each promoter prediction was conducted using 150 base pairs upstream of the predicted sRNA start site. Nucleotide frequency plots were created using the -10 box and -35 box predictions. The web based program WebLogo3 from the University of California at Berkeley was used to generate the sequence logos (Crooks *et al.*, 2004).

TARGET PREDICTION

Target genes for each candidate sRNA were predicted using independent web based programs, TargetRNA2, IntaRNA, and/or CopraRNA (Busch *et al.*, 2008; Tjaden,

2008; Wright *et al.*, 2014). TargetRNA2 searches a genome's annotated features for a statistically significant base pair-binding potential to the queried nucleotide input. The individual base pair model was used throughout target predictions. The program calculates a hybridization score followed by a statistical significance of each potential RNA interaction (Tjaden, 2008). The following parameters were used for each prediction. For statistical significance, the P-value was set at ≤0.05. The program searched 80 nucleotides before the start codon and 20 nucleotides after the start codon. The seed length was 7 consecutive nucleotides and corresponds to the average seed length (6 to 8 nucleotides) for trans-acting sRNAs (Gottesman and Storz, 2011). The filter size, which corresponds to how the program filters out non-target mRNAs, was set at the default value of 400.

Conversely, IntaRNA assesses the query sRNA nucleotide sequence with the selected genome and calculates a combined energy score from the free energy hybridization and interaction sites. This program has less customizable features than TargetRNA2. For these predictions, the parameters were set to default with a minimum number of 7 base pairs in the seed region. For statistical significance, the P-value was set at ≤0.05. Any targets with a P-value >0.05 were discarded from the analysis.

Last, CopraRNA assesses potential targets at the genomic level by integrating phylogenetic information and regulatory networks through a functional enrichment and network analysis. As with the other two programs, default settings were primarily employed as these settings have been optimized to efficiently predict sRNA targets. However, the target region under interrogation was adjusted from a default of 200 nucleotides upstream and 100 nucleotides downstream of the start codon to 200

nucleotides upstream and downstream of the start codon. Targets that were annotated to have a significant interaction ($p<0.05$) in the sRNA-mRNA seed region were retained for further analysis.

CELL CULTURE AND RICKETTSIAL INFECTION

Human dermal microvascular endothelial cells (HMECs) were cultured in MCDB131 medium supplemented with L-glutamine (10 mmol L^{-1}), epidermal growth factor (10 ng mL^{-1}), hydrocortisone (1 μg mL^{-1}), and 10% heat-inactivated fetal bovine serum and grown at 37°C with 5% CO_2 until approximately 80 to 90% confluency (Rydkina *et al.*, 2010). The use of human cell lines in my study was exempt by the University of Texas Medical Branch (UTMB) Institutional Review Board (IRB), but approved by the UTMB Institutional Biosafety Committee (IBC). *Amblyomma americanum* tick cells (AAE2) were grown in L-15B complete medium (pH 7.5) at 34°C to ~90% confluence. Approximately 24h prior to infection, the medium in each flask was replaced with L-15B infection medium (pH 7.5) containing 25 mM sodium bicarbonate and HEPES (Munderloh and Kurtti, 1989). Stocks of *R. prowazekii* strain Breinl were prepared by infecting Vero cells in culture, followed by purification of rickettsiae by differential centrifugation. The titers of infectious stocks were estimated by using a combination of quantitative PCR using primer pair *Rp*877p-*Rp*1258n for citrate synthase gene (*gltA*) and plaque assay (Roux *et al.*, 1997; Rydkina *et al.*, 2010). HMECs were infected with approximately $6x10^4$ pfu of rickettsiae per cm^2 of culture surface area under BSL-3 conditions to achieve an MOI of 5:1 (an average of five or six intracellular rickettsiae per cell and infection of a majority [>80%] of cells (Rydkina *et al.*, 2005; Rydkina *et al.*, 2007; Rydkina *et al.*, 2010). To allow for efficient rickettsial adhesion

and invasion, the cells were incubated for 15 minutes with gentle rocking with an initial infectious inoculum containing *R. prowazekii* in culture medium, prior to replacement with fresh medium. Infected cells were then incubated at 37°C with 5% CO_2 until processing for the isolation of DNA and RNA.

RNA ISOLATION

Depending on the downstream application, total RNA was isolated using several methods and at different time points. For Aim #1, total RNA was isolated at 1.5, 3, 6, 12, 24, 48, and 72 h post-infection using Tri-Reagent® (Molecular Research Center). RNA was extracted using the Direct-zol™ RNA MiniPrep kit (Zymo Research). Column DNaseI treatment (Zymo Research) was performed on all RNA samples to eliminate contaminating genomic DNA. For Aim #2, HMECs infected with *R. prowazekii* as described above were processed for total RNA isolation using Tri-Reagent (Molecular Research Center) at 3 and 24 h post-infection. Total RNA was extracted using the Tri-Reagent® phenol:chloroform method as described by the manufacturer. The RNA samples were treated with DNaseI (Zymo Research) and subjected to the MICROBEnrich Kit (Ambion). The RNAs from the previous step was further purified to eliminate DNaseI protein via precipitation with 100% v/v ethanol prior to the determination of concentration in final preparations. All RNA concentrations were determined using a MultiSkan GO Microplate Spectrophotometer (ThermoScientific).

RNA SEQUENCING

In order to sequence the rickettsial transcriptome, HMECs were infected with *R. prowazekii* strain Breinl, and total RNA was isolated at 3 and 24 h post-infection using Tri-Reagent (Molecular Research Center). Samples were treated with DNaseI (Zymo Research) to eliminate contaminating genomic DNA and subjected to the Dynabeads® Oligo $(dT)_{25}$ (ThermoFisher Scientific) to remove interfering eukaryotic mRNAs. Ribosomal RNA was then removed using the Ribo-Zero kit (Epicentre). RNA was quantified using the MultiSkan Go (ThermoScientific) and analyzed using the Agilent 2100 Bioanalyzer (Agilent Technologies). For each experimental condition, two independent cDNA libraries were created and sequenced on the HiSeq 1500 (Illumina) located in the Next Generation Sequencing Core, University of Texas Medical Branch at Galveston. Enriched RNA used for cDNA synthesis was not size selected, and strand-specific sequencing was performed. Each library consisted of 50bp long subsequences (reads) in a FASTQ format. Each read was assessed for its quality. Any base with a PHRED score of 15 or below was excluded. The first 14 bases of the read were trimmed and the remaining 36 bases were used for analysis. All reads mapping to human genome version GRCh38/hg38 were excluded from the analysis. The remaining rickettsial transcripts were then mapped to *R. prowazekii* strain Breinl genome (NC_020993) allowing up to two base mismatches using Bowtie2 (Langmead and Salzberg, 2012). For each prediction, the average read coverage for each nucleotide was normalized to the length of the predicted sRNA. The same was computed for 50 nucleotides up- and downstream of each prediction. The Mean Expression Value (MEV) was calculated by computing the ratio between the predicted sRNA and the flanking 50 nucleotides

(Raghavan *et al.*, 2011; Warrier *et al.*, 2014). An MEV cutoff value of ≥1.5 was used throughout this work.

For Aim #3, HMECs and AAE2 samples were divided, and one pool respectively underwent Terminator 5'-Phosphate-Dependent Exonuclease (TEX) (Epicentre) treatment in order to eliminate processed RNAs that have a 5' monophosphate. Specifically following the manufacturer's directions, the remaining RNA includes unprocessed, full-length transcripts. Independent cDNA libraries were created using the TruSeq RNA Sample Prep Kit (Illumina) as per manufacturer's directions for each of the non-size-selected RNA samples. Strand-specific sequencing was carried out on an Illumina HiSeq 1500 instrument at my institutional next generation sequencing Core facility. The sequencing libraries comprised 100 bp long reads in a FASTQ format. The quality of each read was assessed, and any base with a PHRED score of 15 or below was excluded from the analysis. The first 14 bases of each read were trimmed, and the remaining 86 bp long reads were then mapped onto the *R. prowazekii* Breinl genome (NC_020993) allowing up to four base mismatches using Bowtie2 (Langmead and Salzberg, 2012). Reads that did not map to the *R. prowazekii* Breinl genome or that mapped more than once to the genome were discarded from analysis.

REVERSE TRANSCRIPTASE PCR

One microgram (1 μg) of DNase I-treated total RNA was reverse transcribed using SuperScript® VILO cDNA Synthesis Kit (Life Technologies) with random hexamers following the manufacturer's instructions. Quantitative RT-PCR was performed on StepOnePlus™ Real-Time PCR System (Applied Biosystems) using

primers designed by Primer Express 3.0.1 (Applied Biosystems). For the TaqMan® Assay, each 20 µL reaction contained 1X TaqMan® Universal PCR Master Mix (Life Technologies), 250 nM forward primer, 250 nM reverse primer, 250 nM TaqMan probe, and 1.1 ng/µL of cDNA. Cycler conditions were: stage 1 at 50ºC for 2 minutes, stage 2 at 95ºC for 10 minutes, stage 3 (40 cycles) at 95ºC for 15 seconds and 60ºC for 60 seconds. Each TaqMan® technical replicate was performed in triplicate using five biological replicates. Primers are listed in Table 3.

Reverse transcriptase PCRs were performed using a Phusion® High-Fidelity PCR Kit (New England BioLabs) in Aim #1. Each 20 µL reaction contained a final concentration of 1X Phusion HF Buffer, 0.2 µM dNTPs, 0.5 µM forward primer, 0.5 µM reverse primer, 100 ng cDNA template, and 0.4 units of Phusion DNA polymerase. Thermal cycler conditions were: stage 1 at 98ºC for 30 seconds, stage 2 (35 cycles) at 98ºC for 15 seconds, 60ºC for 30 seconds, and 72ºC for 30 seconds, and stage 3 at 72ºC for 10 minutes. Samples were separated on a 2% agarose gel, stained with ethidium bromide, and imaged on ChemiDoc MP imaging system (Bio-Rad). Primers are listed in Table 3.

For Aims #2 and #3, reverse transcriptase PCRs were performed using GoTaq® Green Master Mix Kit (Promega) in order to optimize work flow. Each 25 µL reaction contained a final concentration of 1X GoTaq® Master Mix (contains DNA polymerase and dNTPs), 0.5 µM forward primer, 0.5 µM reverse primer, and 100 ng cDNA template. Thermal cycler conditions were: stage 1 at 95ºC for 5 minutes, stage 2 (35 cycles) at 95ºC for 30 seconds, 60ºC for 30 seconds, and 72ºC for 30 seconds, and stage 3 at 72ºC for 10

minutes. Samples were separated on a 1.5% agarose gel, stained with ethidium bromide, and imaged on ChemiDoc MP imaging system (Bio-Rad). Primers are listed in Table 5.

NORTHERN BLOTTING

Northern blot analysis was carried out using the NorthernMax kit (Ambion) following the manufacturer's instructions. Enriched bacterial RNA (15 μg per lane) was loaded onto a 1.5% agarose-formaldehyde gel, separated by electrophoresis at 90V, and then transferred onto a Zeta-Probe Blotting Membrane (Bio-Rad). Membranes were cross-linked using a UV Stratalinker 1800 (Stratagene). A PCR template with a T7 promoter on the anti-sense strand for the sRNA under investigation was created using GoTaq® Green DNA Polymerase (Promega). With the PCR template, strand-specific, [α-^{32}P]UTP-labeled RNA probes were synthesized through *in vitro* transcription with the MAXIscript® kit (Ambion) (Table 5). Each RNA probe was treated with DNase for 15 minutes at 37°C as per manufacturer's directions to remove the original PCR template. Unincorporated nucleotides were removed using Illustra MicroSpin G-25 Columns (GE Healthcare). Membranes were hybridized overnight at 50ºC to 65ºC depending on the probe sequence, washed thoroughly using standard Northern wash solutions, and then exposed to autoradiography film.

RNA LIGASE-MEDIATED RAPID AMPLIFICATION OF CDNA ENDS (RLM-RACE)

The 5' sRNA sequence was determined using FirstChoice® RLM-RACE kit (Ambion) according to the manufacturer's instruction manual. Ten micrograms of

DNase-treated, enriched RNA was incubated with tobacco alkaline phosphatase (TAP) for 1 h at 37ºC. The 5' adaptor sequence was then ligated to TAP-treated RNA at 37ºC for 1 h. Reverse transcription reactions were carried out using random decamers at 42ºC for 1 h. Nested PCR was conducted with necessary modifications to the manufacturer's directions. In order to optimize the cycling conditions yielding consistent amplification, gradient PCR reactions were performed using both the outer and inner primer pairs. Thus, optimal conditions yielding the cleanest and strongest product for each sRNA were employed in all assays. Primers are listed in Table 4. PCR products were cloned into the pGEM-T Easy vector (Promega). Sanger Sequencing was conducted at the institutional Molecular Genomics Core.

ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)

Validation of *Rp*_sR60 and target interactions were through the employment of electrophoretic mobility shift assay (EMSA) following standard protocols (Morita *et al.*, 2012). The DNA templates for *Rp*_sR60 and its target predictions were amplified through PCR using GoTaq® Green Master Mix Kit (Promega) as instructed. Each template incorporated a T7 promoter attached to the amplifying primer in order for *in vitro* transcription using MAXIscript® kit (Ambion) (Table 6). The RNA transcripts for *Rp*_sR60 were radiolabeled with [α-^{32}P] UTP during *in vitro* transcription. After which, the transcripts were purified using Sephadex Microspin Columns (GE Healthcare). *Rp*_sR60 radiolabeled transcripts and mRNA unlabeled transcripts were mixed at a 2:1 molar ratio and incubated at 70°C for 5 minutes followed immediately by 30 minute incubation at 30°C or 37°C. As a control, unlabeled *Rp*_sR60 transcripts were incubated

as a cold competition. Samples were then resolved across a 4% polyacrylamide native gel. Immediately afterwards, gels were dried at 80°C for 55 minutes onto Whatman 3 mm filter paper using a gel dryer. Gels were then exposed to autoradiographic film at varying exposure times.

Table 2. Rickettsial species.

Rickettsia species	RefSeq
Ancestral Group	
R. bellii strain OSU 85-389	NC_009883
R. bellii strain RML 369-C	NC_007940
R. canadensis strain McKiel	NC_009879
Typhus Group	
R. prowazekii strain Madrid E	NC_00963
R. prowazekii strain Breinl	NC_020993
R. typhi strain Wilmington	NC_006142
Transitional Group	
R. akari strain Hartford	NC_009881
R. felis strain URRWXCal2	NC_007109
Spotted Fever	
R. africae strain ESF-5	NC_012633
R. conorii strain Malish 7	NC_003103
R. heilongjiangensis strain 054	NC_015866
R. japonica strain YH	NC_016050
R. massiliae strain MTU5	NC_009900
R. peacockii strain Rustic	NC_012730
R. rickettsii strain Sheila Smith	NC_009882
R. rickettsii strain Iowa	NC_010263
Plasmids	
R. africae strain ESF-5 plasmid pRAF	NC_012634
R. felis strain URRWXCal2 plasmid pRF	NC_007110
R. massiliae strain MTU5 plasmid pRMA	NC_009897
R. peacockii strain Rustic plasmid pRPR	NC_012732

List of *Rickettsia* species and strains used throughout this work. The list has been categorized into the rickettsial groups with the RefSeq ID listed in the next column.

Table 3. PCR Primers.

Gene Name	Strand	Sequence (5' → 3')
6S	Forward	GCTTGGTGCGTGAGCTTAATATT
6S	Reverse	TGAATGAGTATTAGACAGAGACGATGTC
6S	TaqMan	ACCTAGGACTACCACTAAC
16S	Forward	TCGCTAGTAATCGCGGATCA
16S	Reverse	TGTACAAGGCCCGAGAACGT
16S	TaqMan	CATGCCGCGGTGAA
#1	Forward	CTGTCGCTTTTGCCACTATCAT
#1	Reverse	TGACCTTTAGTGCGATTAGTCACAA
#2	Forward	ATCTTGCTCTTGGTGGGGTTA
#2	Reverse	AGGGTGAAAAGGTGTGGTAAGAG
#5	Forward	CACTTTAGTGTAGTCATAATTTCTCAA
#5	Reverse	TGTTAAAAAGTGTACGATCAGATAAT
#9	Forward	TCAAGGCAGAACATAACATAAGCTAGA
#9	Reverse	AATCACCATTGTGAGCATCAATAAG
#10	Forward	AAGCTCTGAATCGTGAATCTCTTAAGT
#10	Reverse	GCTGCCATAATTACCTCCTCAATT
#24	Forward	GCACAAAACTTTAAGATGCAGAAAAA
#24	Reverse	TCCTGCTGTTATGTCAGTACTTTGAA
#25	Forward	TGTTTCCTCCCCCATGACTT
#25	Reverse	TTCCTACATTTGCAAGGATTACT
AAE4_sRNA	Forward	TGTTAAGAAGTTAGTGCCAAAACAA
AAE4_sRNA	Reverse	ACCTTAAGAGCTTAGCAAAAAGCTA
AAE6_sRNA	Forward	TCACAGAGTTAATATGCAGGTCTCT
AAE6_sRNA	Reverse	GCGCAGTTTAAATGTGATTATATGC
AAE15_sRNA	Forward	AAATGCGACTTCACAATAACTGA
AAE15_sRNA	Reverse	TACAAAATTGACTATCCCACATGGT
AAE18_sRNA	Forward	CGGTCCTATTAGGCTTGAAATTTGT
AAE18_sRNA	Reverse	AGCATCACTCGTAGAAAGTACAGA
AAE36-sRNA	Forward	ACAGAAAAACTATTGCATAGGAAGC
AAE36_sRNA	Reverse	ATCCAATGCTGATGTTCGCT
AAE63_sRNA	Forward	TAGCGTTACTCCCCTGACACT
AAE63_sRNA	Reverse	GTGATACTGAGACGCCATTAGTGA
AAE72_sRNA	Forward	TGTTTTCATACACGATCTAGCACC
AAE72_sRNA	Reverse	TGCCACTGCATCTACAAATCTT
AAE96_sRNA	Forward	TCGAGGTAATGTGCTGATTGTT
AAE96_sRNA	Reverse	ACTGGTGTTGTGCTAGTGCTTA

Primers utilized for the RT-PCR analysis of *R. prowazekii* novel sRNAs.

Table 4. RLM-RACE primers for transcription start site.

*Rp*_sR	Direction	Sequence (5' →3')
67	5' Outer Primer	AATTGTTTCCTCCCCCATGACT
67	5' Inner Primer	CCCCATGACTTCATACTGTGCTT
17	5' Outer Primer	GGCAAAATATGTTACTCATTGCTG
17	5' Inner Primer	TACATGGGTTTCACACTGCCT
34	5' Outer Primer	ACTTTGTGGTCAAAAGTTGCTAA
34	5' Inner Primer	ACCGTCAGAGGCTCAAAAAGT
60	5' Outer Primer	GCATACTAAAAAAGTTCAAATTCAG
60	5' Inner Primer	ATTACTGCATCTGCACAAACG
47	5' Outer Primer	TCGTTAAAAGCCCCAAAAGTACC
47	5' Inner Primer	GCTGGAGCTTTGCAAGATCATA
6S	5' Outer Primer	TATTAGCTGCTTGGTTCGTGAG
6S	5' Inner Primer	GCTTGGTGCGTGAGCTTAATATT

RLM-RACE primers utilized in this study for 5' end sequencing of novel *R. prowazekii* sRNAs.

Table 5. Northern blot primers.

*Rp*_sR	Direction	Sequence (5' → 3')
67	Forward	GAAAT***TAATACGACTCACTATAGGG***ATTGTTTCCTCCCCCATGAC
67	Reverse	TTCCTACATTTGCAAGGATTACT
17	Forward	GAAATTAATACGACTCACTATAGGGAGGCAGTGTGAAACCCATGT
17	Reverse	TGTTACTCATTGCTGCTAGCTAAA
34	Forward	TCAGGTTTTGGAGCAATGGGG
34	Reverse	GAAAT***TAATACGACTCACTATAGGG***TCAAGGATTGGATAAAGGACAAACT
60	Forward	GATTTGACTGCAGAGAGTATTTTGA
60	Reverse	GAAAT***TAATACGACTCACTATAGGG***ACTGCATCTGCACAAACGAT
47	Forward	AGCCTGAACTGATCCTTTACCA
47	Reverse	GAAAT***TAATACGACTCACTATAGGG***ACGGTAATTCTGCATCTGCT
6S	Forward	GAAAT***TAATACGACTCACTATAGGG***TGGCCGTTACTTAAAATCCTTCA
6S	Reverse	TGCTTGGTTCGTGAGCTTAATATT

Primers utilized for the generation of strand-specific probes for Northern blot analysis of *R. prowazekii* novel sRNAs. Underlined is the T7 promoter sequence used for in vitro transcription.

Table 6. EMSA primers.

Gene	Strand	Sequence (5'→3')
*Rp*_sR60	Forward	GAAAT***TAATACGACTCACTATAGG***GTGACTGCAGAGAGTATTTTGAGT
*Rp*_sR60	Reverse	GCCAAAATGCATTGTACTCTCT
H375_8740	Forward	GAAAT***TAATACGACTCACTATAGG***AAAATTAACAAAGCGTCATGCAAT
H375_8740	Reverse	TGACTCCGTACCTTAGGTCCAT
H375_7670	Forward	GAAAT***TAATACGACTCACTATAGG***GAGATTAATGGAAGAAGTTGAGCATA
H375_7670	Reverse	GCTGTTTCTTTATTTCTTCAAGTTCA
H375_2860	Forward	GAAAT***TAATACGACTCACTATAGG***CGCTTTGCTCCATCACCAAC
H375_2860	Reverse	TGTGGTTTTCTGCCCTGTTCT
H375_7680	Forward	GAAAT***TAATACGACTCACTATAGG***GAGATTAATGGAAGAAGTTGAGCAT
H375_7680	Reverse	GCTGTTTCTTTATTTCTTCAAGTTCA
H375_6080	Forward	GAAAT***TAATACGACTCACTATAGG***AGGTACAATGGTAGGGCTCTT
H375_6080	Reverse	ACTATTTCAGCACTCACTTGT
H375_8760	Forward	GAAAT***TAATACGACTCACTATAGG***TCACAAGAAGCGGTTAGCTT
H375_8760	Reverse	TCAACTTGCACTAGCTCTCT
H375_420	Forward	GAAAT***TAATACGACTCACTATAGG***TACATTGTTTTAATCTTCCCCACCAT
H375_420	Reverse	AATGCCTCAACTTCAGAATTAGTG

Primers utilized for the generation of strand-specific probes for electrophoretic mobility shift assay (EMSA) analysis of *R. prowazekii* novel sRNAs. Underlined is the T7 promoter sequence used for in vitro transcription.

Chapter 3 Prediction of Small RNAs in the Genus *Rickettsia*[2]

INTRODUCTION

Due to the historic importance of rickettsial diseases, their global distribution and associated morbidity or mortality, and potential implications in bioterrorism, the genome of *R. prowazekii* was the first to be sequenced and published (Andersson *et al.*, 1998). Unlike other intracellular bacteria, whose genomes have very high coding densities, *R. prowazekii* was found to have 24% non-coding DNA (Andersson *et al.*, 1998; Holste *et al.*, 2000). Such a large amount of non-coding DNA was thought to be the consequence of genomic reduction and pseudogenization due to the loss or degradation of genes involved in several mechanisms leading to the obligate intracellular lifestyle of this pathogen. A number of other rickettsial genomes, including those of *R. rickettsii, R. conorii, R. typhi,* and other notable species have since been sequenced and either published or made available in biomedical databases (Ogata *et al.*, 2001; McLeod *et al.*, 2004; Gillespie *et al.*, 2012). However, the presence of small, non-coding RNAs in different *Rickettsia* species still remains undetermined. With an aim to address this important knowledge gap, I used SIPHT/sRNAPredict2 to identify candidate novel sRNAs within the intergenic regions of all four rickettsial groups, leading to the prediction of a total of 1,785 novel sRNAs within 16 different strains representing 13 rickettsial species. I further analyzed the predicted sRNAs in *R. prowazekii* strain Breinl using other bioinformatics tools and experimentally validated their expression using RT-

---

[2] The work mentioned in this chapter is based on the work published: **Schroeder, C. L**., H. P. Narra, M. Rojas, A. Sahni, J. Patel, K. Khanipov, et al. (2015). Bacterial small RNAs in the Genus *Rickettsia*. *BMC Genomics* 16(1)**,** 1075. doi: 10.1186/s12864-015-2293-7.

PCR and deep sequencing approaches. In tandem, these analyses constitute the very first evidence documenting the presence of sRNAs in rickettsial genomes.

RESULTS

Bioinformatics Prediction of Small RNAs

Ready availability of complete genome sequences render computational approaches a widely acceptable first step for identification of sRNAs (Livny *et al.*, 2006; Livny *et al.*, 2008; Lu *et al.*, 2011). Such bioinformatics approaches search the intergenic regions (IGRs) of a bacterial genome for specific sRNA features. In general, the strategy involves an in-depth search of IGRs for the presence of Rho-independent terminators, promoters, and transcription factor binding sites, followed by the analysis of their secondary structure and comparison of such IGRs with closely related species (Livny *et al.*, 2008; Sharma *et al.*, 2010; Sharma *et al.*, 2014). In this study, I employed the web-based program SIPHT to examine the genomes of 16 rickettsial strains, which represent a total of 13 species spanning all four rickettsial groups. Four known plasmids, representing three from the spotted fever group strains and one from a transitional group strain, were also included. The *R. felis* pRFdelta plasmid (NC_007111.1) was excluded from the analysis as it was found to be an artifact from genome assembly (Gillespie *et al.*, 2007). I primarily employed recommended default settings to identify most sRNA features. To perform a more stringent search, however, I chose to decrease the Expectation Value (E value) from the default $5e^{-3}$ to $1e^{-15}$ to minimize false positives. As a result, I identified a total of 1,785 candidate rickettsial sRNAs (Additional File 1). On average, I predicted 74 candidates per ancestral strain, 21 candidates per typhus strain,

152 candidates per transitional strain, and 158 candidates per spotted fever strain. Table 7 categorizes the number of predictions by nucleotide size and rickettsial strain. Of the four plasmids examined, *R. peacockii* plasmid RPR had a single sRNA prediction.

Computational Analysis of sRNA Predictions

Based on the predictive analysis suggesting sRNAs in all rickettsial groups, I set out to analyze each of the candidate sRNAs. In order to examine a common set of bacterial sRNAs within *Rickettsia* species, I first investigated five well-known bacterial sRNAs, namely 6S RNA (*ssrS*), α-tmRNA (*ssrA*), RNaseP_bact_a, rpsL_ricks, and 4.5S RNA (*ffs*), and confirmed their presence in *R. prowazekii*. Using BLAST with an E-value cut-off of $1e^{-5}$, I compared the prediction for each strain against others included in the study to find shared sRNA candidates (Altschul *et al.*, 1990). As expected, I noted that rickettsial strains closely related to each other phylogenetically had a greater number of shared sRNA candidates when compared to those that are distantly related (Table 8). For example, *R. rickettsii* strain Sheila Smith (virulent) and strain Iowa (avirulent), which share 96.6% homology (Ellison *et al.*, 2008; Clark *et al.*, 2015), had 115 sRNAs in addition to ten other sRNA predictions in Iowa, and 20 present only in Sheila Smith.

I next employed the web-based program, BPROM, to determine promoter motifs for predicted sRNAs in the spotted fever and typhus groups of rickettsiae, since they represent the most prominent human pathogens. While it is still undetermined whether the ancestral group causes human disease, the transitional group species are established pathogens, but they account for a small fraction of reported rickettsiosis cases. Using all sRNA predictions within the typhus group (3 strains) and the spotted fever group (8

strains), I searched 150 nucleotides upstream of the predicted sRNA start site for the -10 and -35 promoter motifs because nearly 80% of known σ^{70} promoters in *E. coli,* considered to be a model organism, fall within 150 bp of the transcription start site (Huerta and Collado-Vides, 2003). While BPROM successfully predicted the -10 and -35 promoter sites for all candidate typhus group sRNAs, it was unable to predict a promoter site for one sRNA candidate (#132) belonging to *R. rickettsii* strain Sheila Smith. Using the data obtained from the BPROM software, the average distance for the predicted -10 motif and -35 motif for both spotted fever and typhus groups of rickettsiae was calculated. For the typhus group, -10 and -35 motifs were an average of 67 and 88 (Stan. Dev. = ±22) nucleotides upstream of the sRNA start site, respectively. The spotted fever group had similar nucleotide distances at 70 and 91 (Stan. Dev = ±22) nucleotides upstream. The average distance between the -10 and -35 motifs was 21 nucleotides for both groups, slightly longer than the reported 17±1 nucleotide distance optimal for *Escherichia coli* genes (Harley and Reynolds, 1987; Mitchell *et al.*, 2003). Typhus group and spotted fever nucleotide frequencies for the -10 and -35 motifs were plotted using WebLogo3 (Figure 4).

The known consensus sequences of *E. coli* ORFs for the -10 motif and -35 motif are TATAAT (with each nucleotide probability at 82%, 89%, 52%, 59%, 49%, 89%) and TTGACA (with each nucleotide probability at 69%, 79%, 61%, 56%, 54%, 54%), respectively (Harley and Reynolds, 1987). Both rickettsial groups favored a -10 motif similar to the consensus sequence. Interestingly, the -35 motif differed between the groups, as well as from the *E. coli* consensus sequence, at the fifth nucleotide. In *E. coli*, this position is cytosine in 54% of tested sequences. However, it is adenosine

(approximately 41%) or thymine (approximately 40%) in spotted fever group and thymine (approximately 35%) or adenosine (approximately 30%) in typhus group. In addition, the second nucleotide position in typhus group is most conserved with a thymine in nearly 100% of predicted sites, while it is approximately 90% for spotted fever group. This is in contrast to *E. coli*, which has a thymine with 79% probability at the same nucleotide position.

In an attempt to explain the differences in sRNA predictions between Sheila Smith and Iowa strains of *R. rickettsii*, I performed a comparative analysis by mapping the sRNAs present only in one of the strain but absent in another. Accordingly, 20 predictions from Sheila Smith strain and their corresponding 150bp up- and downstream sequences to the strain Iowa genome were compared. All "prediction±150 bp" sequences were >99% identical with the exception of two (#23 and #71) sRNAs. For instance, sRNA candidate #71 had a 20 bp sequence absent from the predicted sRNA sequence in the Iowa genome (Figure 5). The same analysis was conducted for the 10 predictions present only in *R. rickettsii* strain Iowa but absent in strain Sheila Smith. Again, all but two "prediction±150 bp" sequences were nearly identical (>99%). In this case, prediction #118 for strain Iowa had a 46 bp sequence that was absent in strain Sheila Smith (Figure 6). Also, an ORF was annotated in corresponding genomic regions in strain Sheila Smith for nearly 30% of sRNAs predicted only in strain Iowa, while two sRNAs had SNPs and indels in the Sheila Smith sequences potentially leading to altered thermostability of secondary structures. Similar observations were made for sRNAs present only in Sheila Smith but absent in Iowa strain. Since SIPHT relies on the conserved intergenic regions, secondary structures, presence of promoters, and terminator

sequences, it is likely that the sRNAs predicted only in one strain, but not the other, result from these stringent criteria.

Candidate sRNA Target Identification

Although spotted fever rickettsiae cause disease worldwide, I chose to initially focus on *R. prowazekii* due to the high public health threat. In addition, due to bioweapon testing during World War II and the development of antibiotic-resistant strains during the Cold War, *R. prowazekii* remains on the list of select agents with potential for bioterrorism (Walker, 2003). To identify potential mRNA targets for predicted sRNAs within the *R. prowazekii* genome, two independent web-based programs, TargetRNA2 and CopraRNA, were chosen to predict sRNA:mRNA interactions by assessing the base pairing potential based on a Smith-Waterman dynamic and conservation profile, respectively (Tjaden, 2008; Wright *et al.*, 2013; Wright *et al.*, 2014). The search parameters were set to default and a p-value threshold of ≤0.05. A total of 393 potential targets were identified by TargetRNA2, whereas CopraRNA predicted 1154 protein coding genes to be regulated by sRNAs. A detailed comparative analysis revealed that 16 sRNA candidates had common target genes predicted by both programs and the remaining 10 candidates had independent predictions (Additional File 2). Two sRNA candidates (#6 and #22) had the highest number of 6 common targets predicted by TargetRNA2 and CopraRNA, in contrast to only one commonly predicted target for candidates #12 and #19. In summary, a total of 51 target genes were predicted by both programs, of which only 9 were categorized as hypothetical proteins. Of note, target genes such as *virB10, ftsL, ftsQ, secA, ruvB,* and 190kDa antigen were commonly

predicted, indicating the potential role of post-transcriptional regulatory mechanisms in type IV secretion, cell division, and DNA repair. A majority of the common targets had functional roles. Table 9 lists the number of predicted target transcripts for each predicted sRNA. Interestingly, TargetRNA2 failed to predict targets for candidate #7, but CopraRNA predicted 53 target genes (Table 9). Nevertheless, if the lack of target prediction by TargetRNA2 holds true for candidate #7, this may be due to a sRNA:protein interaction, as is the case with 6S RNA, or it may simply represent a degraded ORF with an active promoter and terminator. Using this information, the predicted mRNA targets were parsed based on their respective protein functions. Eight categories were selected, including a category for 'other' (annotated ORFs with known function, but not categorized into a separate class based on function) and 'hypothetical proteins' (annotated ORFs with unknown/uncharacterized function), and separated the 393 predictions into different categories. The categories included cell division, cell wall, metabolism, ribosomal functions, virulence, type IV secretion system, transport proteins, and phagosomal escape (Table 10). My TargetRNA2 and CopraRNA results demonstrate that the majority of known targets for sRNAs are involved in metabolism (71 *vs.* 197), ribosomal functions (51 *vs.* 129), and cell division (44 *vs.* 30). However, both the programs predicted a large number of target genes potentially regulated by sRNAs that were categorized as 'other' (90 *vs.* 370) and 'hypothetical proteins' (73 *vs.* 232), respectively (Table 10). While TargetRNA2 uses Smith-Waterman dynamic based base pairing, CopraRNA predictions are largely based on conservation between different genomes, possibly resulting in the differences in the number of predicted targets.

RNA Sequencing

Next, I set out to confirm the expression of rickettsial sRNAs during infection of cultured human microvascular endothelial cells (HMECs) with *R. prowazekii*. HMECs were infected with *R. prowazekii* strain Breinl and total RNA extracted at 3 and 24 h post-infection. My rationale for choosing these durations was to allow ample time for rickettsial entry and establishment of infection within the host cells and sufficient time for at least two replication cycles keeping in mind that replication time for intracellular rickettsiae ranges from 9 to 11h. After removal of rRNAs and eukaryotic mRNA, the enriched bacterial RNA was reverse transcribed into cDNA libraries and subjected to next generation sequencing using the Illumina HiSeq^{TM} 1500 system. The resulting RNA reads were mapped onto the *R. prowazekii* strain Breinl (NC_020993) genome. My deep sequencing resulted in approximately 42.6 to 46.2 million total reads for RNA isolated at 3 h and 27.4 to 28.6 million total reads at 24 h post-infection. Out of these, an average of 1.4 and 2.8 million reads mapped to the *R. prowazekii* genome at 3 h and 24 h, respectively. *Rickettsia* species include obligately intracellular bacteria with fastidious growth requirements in a host cell and cannot yet be cultured in a cell-free environment. Recently, it has been reported that intracellular organisms such as *Rickettsia* represent only 5% of the extracted total RNA, while the remaining 95% belongs to the eukaryotic host. Out of approximately 5% bacterial total RNA, 95% is composed of ribosomal and transfer RNA, while the remaining 5% of the transcripts correspond to bacterial mRNA and sRNA, yielding a ratio of ~1:400 bacterial mRNA and sRNA in total RNA extracted during the infection (Westermann *et al.*, 2012). Although microbe enrichment is aimed at removing most of the polyadenylated eukaryotic transcripts and ribosomal RNAs, the

process often accomplishes only limited removal of other interfering eukaryotic RNAs such tRNAs, noncoding RNAs, and mitochondrial RNA. Furthermore, high abundance of rRNAs in the host cells also interferes with the efficacy of their removal from the sample preparations. Supporting my results, a recent study has reported that only 2-5% of the total reads mapped to the intracellular bacterial genomes despite enrichment of the total RNA (Westermann *et al.*, 2012). By analyzing the sequencing data at the genome locations predicted by SIPHT, twelve out of 26 predicted sRNAs had a Mean Expression Value (MEV) that was ≥1.5 times compared to their respective 50 nucleotide upstream and downstream flanking regions (Table 11). As expected, all five well known sRNAs, namely 6S RNA, α-tmRNA, RNaseP_bact_a, rpsL_ricks, and 4.5S RNA were found to be expressed *in vitro* and exhibited an MEV of >1.5. The read coverage plots of 6S RNA, RNaseP_bact_a and α-tmRNA are presented in Figure 7.

Validation of sRNA Predictions via RT-PCR.

Prior to validating the predicted sRNAs, I decided to investigate the expression of 6S RNA within *R. prowazekii.* The underlying rationale for choosing 6S RNA to begin with was its particularly high abundance in *E. coli,* which can reach ~10,000 copies during late stationary phase (Wassarman and Storz, 2000). Using 16S rRNA as the endogenous control and infection for 1.5 h as baseline, I found a significant ($p<0.01$) increase in its expression from 6 to 72 h post-infection (n=5) using TaqMan-based real-time RT-PCR (Figure 8). After confirming expression of 6S RNA, nine sRNA candidates (#1, #2, #5, #9, #10, #11, #21 #24, and #25) were chosen to verify their expression in *R. prowazekii* str. Breinl during infection of HMECs. These were chosen

based on their location within the genome, orientation comparative to the neighboring genes, and potential mRNA targets (Figure 9). Candidates #11 and #21 were not detected using RT-PCR. However, the remaining seven candidates were detected using RT-PCR (n=3) with an amplicon near the expected size.

Figure 10 shows a representative agarose gel for candidates #1, #5, #9, #10, #24, and #25. Upon cross-reference with other *R. prowazekii* genomes that included strains Madrid E, Dachau, BuV67, Katsinyian, Chernikova, RpGvF24, GvV257, and Rp22, it was found that candidate #2 was anti-sense to the gene *rnpB* (RNaseP_bact_a) annotated only in *R. prowazekii* strain Rp22 (NC_017560). Therefore, any amplification is likely the result of *rnpB* expression. The remaining predictions demonstrated no association with any other annotated open-reading frames. To further confirm this observation, each sRNA sequence was examined for its ability to code for a protein. Using the ExPASy Translate Tool (Swiss Institute of Bioinformatics), all six possible translation initiation positions (3 each on 5' and 3' strands) were assessed to be devoid of protein coding capacity, yielding evidence that these are indeed small non-coding RNAs expressed during rickettsial infection of host endothelium.

DISCUSSION

In this study, I did a genome wide computational analysis to identify novel sRNAs within the genus *Rickettsia.* I identified 1,785 sRNAs in 16 rickettsial strains belonging to 13 different species and spanning across all rickettsial groups. To further confirm my sRNA predictions, I have validated the expression of six predicted trans-acting sRNAs in *R. prowazekii* strain Breinl using high throughput sequencing and RT-PCR approaches. Since the initial discovery of sRNAs in 1960s, *E. coli* has been shown

to harbor nearly 80 to 100 small RNAs, while *Salmonella enterica* serovar Typhimurium genome encodes for ~140 small RNAs (Gottesman and Storz, 2011; Shinhara *et al.*, 2011; Kroger *et al.*, 2012). Abundant evidence now demonstrates the ubiquitous nature of sRNAs in bacterial genomes and implicates them in playing an important role in virulence, quorum sensing, survival, plasmid expression, and primary and secondary metabolism in addition to several other housekeeping functions (Wagner and Simons, 1994; Wassarman, 2002; Lenz *et al.*, 2004; Guillier and Gottesman, 2006; Lee and Groisman, 2010; Mraheil *et al.*, 2011; Bobrovskyy *et al.*, 2015). In *Vibrio cholerae* and *V. harveyi,* quorum-sensing genes *hapR* and *luxR* are under the regulatory control of four and five sRNAs, respectively (Lenz *et al.*, 2004). Furthermore, the deletion of three sRNAs in *Listeria monocytogenes* results in an attenuated phenotype in mouse models, and the mutant strain is unable to grow in murine macrophages (Mraheil *et al.*, 2011). Similarly, *rli38* knockout mutants of *Listeria* were found to be attenuated in orally-infected mice, suggesting a role in the pathogen's virulence (Toledo-Arana *et al.*, 2009), and the deletion of *lhrA* in *L. monocytogenes* was capable of altering the expression of over 300 genes (Nielsen *et al.*, 2011). In *Salmonella enterica,* the AmgR small RNA controls the expression of the *mgtCBR* mRNA required for survival in macrophages, and its overexpression leads to decreased virulence in mouse models (Lee and Groisman, 2010). However, studies examining the potential regulatory roles of sRNAs in obligate intracellular bacteria remain rather limited. In addition to a sRNA that regulates *hctA*, 16 *trans*-acting and 25 *cis*-acting sRNAs have been identified in *Chlamydia trachomatis*, an intracellular human pathogen (Grieshaber *et al.*, 2006; Albrecht *et al.*, 2010). More recently, *Coxiella burnetii* and *Buchnera aphidicola* genomes were shown to encode 14

and 140 sRNAs, respectively, and *Coxiella* sRNAs exhibit differential expression at different growth stages (Leroy *et al.*, 2010; Hansen and Degnan, 2014; Warrier *et al.*, 2014). Here, I report on the existence of novel small RNAs in rickettsial genomes and their potential roles as determinants of pathogen virulence, host adaptation, and metabolism.

Several prediction programs using parameters such as comparative genomics, RNA structure, and thermodynamic stability, have been developed and utilized to identify bacterial small RNAs (Gautheret and Lambert, 2001; Rivas and Eddy, 2001; Washietl *et al.*, 2005; Sridhar *et al.*, 2010; Wright *et al.*, 2013). In this study, I have chosen SIPHT to predict trans-acting sRNAs in rickettsial genomes as this prediction tool uses several other well-established and widely used programs to identify potential transcription factor binding sites (Altschul *et al.*, 1997; Liu *et al.*, 2001; Pain *et al.*, 2015); Rho-independent terminators using RNAMotif (Macke *et al.*, 2001), TransTermHP (Kingsford *et al.*, 2007), and FindTerm (Argaman *et al.*, 2001); conserved secondary structures by QRNA (Rivas and Eddy, 2001); and conserved nucleotide sequences by BLASTN 2.0 (Altschul *et al.*, 1997). Further, this program has been widely applied for sRNA predictions in several other bacteria attesting to its potential for accurately predicting bacterial sRNAs, and its web-based availability makes it both user-friendly and easily accessible. Additionally, unlike its counterparts such as eQRNA and RNAz, SIPHT specifically searches for Rho-independent terminators and conserved intergenic structures significantly eliminating the chances of false-positive predictions (Lu *et al.*, 2011). Using SIPHT, I predicted an average of 21, 74, 152 and 158 sRNAs in typhus, ancestral, transitional, and spotted fever groups of *Rickettsia* species,

respectively. To determine if predicted sRNAs have upstream transcription factor binding sites and downstream Rho-independent terminator (two independent criteria used by SIPHT), I further analyzed all *R. prowazekii* sRNAs using BPROM (Solovyey and Salamov, 2010) and TransTermHP (Kingsford *et al.*, 2007). All *R. prowazekii* sRNAs were found to have a predicted upstream σ^{70} promoter and a Rho-independent termination confirming the results retrieved from SIPHT (Figure 4).

My genus-wide global analysis suggests that the repertoire of predicted sRNAs is independent of the size of the respective rickettsial genome. This is exemplified by the presence of 191 sRNAs in the 1.31 Mbp genome of *R. peacockii* (spotted fever), in contrast to only 100 sRNAs in the 1.54 Mbp genome of ancestral *R. bellii*. The number of sRNAs among rickettsiae in different groups, however, tends to directly correlate with their respective genome size, while the average number of sRNAs per Mbp of the genome within a particular group varies depending on the *Rickettsia* species/strain. For example, *R. bellii* and *R. canadensis,* belonging to the ancestral group and carrying the genomes of 1.54 Mbp and 1.15 Mbp, respectively, encode 100 and 47 sRNAs. On the other hand, *R. canadensis* has only 40 sRNAs/Mbp, while *R. bellii* has 65 sRNAs/Mbp, indicating the impact of genomic content and organization on the prediction of sRNAs. Also, although the average length of sRNAs in *R. bellii* and *R. canadensis* is fairly similar (132 *vs* 149, respectively), a detailed analysis of the length of intergenic regions (IGRs) in *R. bellii* and *R. canadensis* revealed that *R. bellii* has ~63% more IGRs ranging from 1-300 bp. It is, therefore, possible that the lower number of sRNAs in *R. canadensis* is due to the differences in the number of IGRs included in the SIPHT analysis. Another notable difference is evident between *R. akari* and *R. felis* in the transitional group.

While both had over 100 predictions, the coding density of sRNAs in *R. akari* (94 sRNAs/Mbp of genome) was 22% lower in comparison to *R. felis* (121 sRNAs/Mbp of genome). *Rickettsia* are generally presumed to exhibit limited horizontal gene transfer (HGT) due to their obligately intracellular lifestyle. However, recent reports document the dynamic nature of their genomes, and transposable elements, palindromic repeats, and horizontally-acquired genes have been identified in several *Rickettsia* species (Gillespie *et al.*, 2007; Gillespie *et al.*, 2008; Sahni *et al.*, 2013). For example, the transposable elements of *R. felis* cause inactivation of genes or integration of foreign DNA resulting in changes to both genomic content and arrangement (Ogata *et al.*, 2005). In this regard, at least 79 genes in *R. felis* have been suggested to be acquired through HGT from other proteobacteria or amoebae (Merhej *et al.*, 2011). Additionally, *R. prowazekii* and *R. typhi*, despite having similar size genomes, encode 26 and 15 sRNAs, respectively. Again, the number of IGRs included in the SIPHT analysis varies between *R. prowazekii* strain Breinl and *R. typhi* strain Wilmington (540 vs 504), which may explain the differences in the total number of predicted sRNAs in these typhus rickettsiae genomes.

Although computational approaches yield convincing evidence for the existence of sRNAs in *Rickettsia*, it is critical to validate the expression of predicted sRNAs during host-pathogen interactions via experimental strategies. In this context, I first attempted to confirm the expression of 6S RNA (*ssrS*), a well-characterized small, noncoding RNA ubiquitously present in most bacterial lineages, including *γ-proteobacteria* and *Bacillales* (Hindley, 1967; Warrier *et al.*, 2014). Although most bacteria encode a single copy of the *ssrS* gene, some bacterial species, including *Bacillus subtilis*, reportedly encode two copies that are differentially expressed depending on the stage of growth (Wassarman,

2007). 6S RNA is most abundantly expressed during late stationary phase, where it interacts with RNA polymerase and regulates σ^{70} function. These data further support possible roles for 6S RNA in long-term survival and nutrient uptake (Trotochaud and Wassarman, 2004; Faucher *et al.*, 2010). Also, 6S RNA is potentially involved in intracellular stress response in *C. burnetii* and *Legionella pneumophila* (Weissenmayer *et al.*, 2011; Warrier *et al.*, 2014). I found a significant increase in the 6S RNA expression from 6 to 72 h post-infection when compared to the basal expression level at 1.5 h. Also, based on my MEV calculations, I observed an increase of 2- and 5-fold in the expression of 6S RNA at 3 and 24 h post-infection, respectively. This is in agreement with earlier findings from other pathogenic bacteria. Specifically, a 2-fold up-regulation in expression at 72 h appears to correspond to a similar increase seen in *C. burnetii* (Warrier *et al.*, 2014). My data further suggest that the highest expression at 72 h post-infection coincides with the intracellular growth kinetics of rickettsiae (Eremeeva *et al.*, 2003). A comprehensive analysis to elucidate its mechanisms of action and regulatory functions in intracellular rickettsiae is warranted and currently in progress.

Because RNA sequencing is a novel and robust methodology that provides valuable insights into the global transcriptome (Ozsolak and Milos, 2011), I subjected total RNA from *R. prowazekii*-infected endothelial cells to validate the presence of sRNAs and their expression during host-pathogen interactions. Since the major focus of this study was the identification and validation of intergenic trans-acting sRNAs, additional information on the *cis*-acting sRNAs expressed by *R. prowazekii* strain Breinl during infection of HMECs was excluded to perform a direct comparative and confirmatory analysis of SIPHT-based sRNA predictions *versus* their expression *in vitro*.

Furthermore, none of the web-based sRNA prediction tools have the ability to identify *cis*-acting sRNAs in genomes. MEV-based identification of sRNAs in bacterial genomes is a widely utilized approach that exploits the expression profile of sRNAs and their respective flanking regions to determine the biogenesis of sRNAs (Raghavan *et al.*, 2011; Moody *et al.*; Warrier *et al.*, 2014). To this end, the reads mapping to each nucleotide of the sRNAs and their respective 50 bp flanking regions were normalized using the total number of reads mapping to the rickettsial genome (excluding those mapping to the rRNAs and tRNAs) and then to their length. MEVs were determined to decipher the expression of sRNAs from potential read-throughs attributed to flanking ORFs due to leaky transcriptional termination in *R. prowazekii* (Woodard and Wood, 2011). Nearly 50% of predicted sRNAs in *R. prowazekii* exhibited an MEV of ≥1.5 when compared to respective flanking regions indicating their biogenesis and expression independent of neighboring genes. The reads' coverage plots for 6S RNA, RNaseP_bact_a, and α-tmRNA clearly demonstrate the independent expression of these sRNAs (Figure 7).

During invasion of epithelial cells and intracellular replication within macrophages, *Salmonella* expresses IsrM RNA encoded in *Salmonella* pathogenicity islands (Gong *et al.*, 2011). *L. monocytogenes* encodes a thermosensor sRNA, which upon encountering human body temperature (37°C) forms an alternative secondary structure and activates adhesins, phagosome escape mechanisms, and other immune-regulating factors (Gripenland *et al.*, 2010). Further assessment of the expression profile of five sRNAs (#5, 9, 10, 24, & 25) showing an MEV of ≥1.5 revealed their steady-state expression of these sRNAs from 1.5 to 72 h post-infection indicating potential roles in pathogenesis. Interestingly, despite an MEV of <1.5, candidate #1 was expressed

between 6 to 72 h and its expression increased over time, most notably after 24 h. Since my in-depth transcriptome analysis was performed at 3 and 24 h post-infection, I hypothesize that the relative expression of candidate #1 is likely inadequate to generate sufficient reads to achieve an MEV greater than 1.5 fold. Alternatively, the sequencing depth may not have been high enough to achieve a 1.5-fold difference. Even though 50% of the predicted sRNAs were either not detected or expressed below the cut-off MEV in my RNA-seq analysis, it is plausible that they are bonafide sRNAs conditionally expressed during other conditions such as stress and host-vector interactions. Previous studies have shown that different environments induce specific sRNAs. Small RNA ryhB, known to down-regulate genes involved in iron storage in *E. coli*, is induced mainly during low iron conditions (Masse and Gottesman, 2002). In *Salmonella enterica* serovar Typhimurium, IsrJ sRNA is induced under low oxygen and magnesium environments and elevated levels of IsrE are observed in iron-responsive environment (Padalon-Brauch *et al.*, 2008). Similarly, *H. pylori* is known to induce the expression of six small RNAs (IsoA1-6) associated with acid stress (Sharma *et al.*, 2010). Alternatively, it is also plausible that SIPHT may have identified degraded ORFs (Andersson *et al.*, 1998; Fournier *et al.*, 2009). Rickettsial genomes are known to evolve by reductive evolution (gene degradation) and transposons are known to play a pivotal role in gene inactivation (Blanc *et al.*, 2007; Gillespie *et al.*, 2007; Gillespie *et al.*, 2008; Ammerman *et al.*, 2009; Gillespie *et al.*, 2012). *R. prowazekii* is known to have pseudogenes potentially resulting from gene inactivation (Andersson *et al.*, 1998). Since SIPHT uses the presence of an upstream promoter and downstream transcriptional terminator as the main criteria for predicting sRNAs, it is possible that some of sRNA

transcripts predicted by SIPHT potentially map to the degrading ORFs, which still retain conserved promoter and terminator regions.

Despite the abundance of sRNAs in all bacterial lineages, little is known about their function and mechanism of action within the bacterial genomes and only a few sRNAs have been assigned functions, to date (Warrier *et al.*, 2014). Using TargetRNA2 and CopraRNA, I have predicted the target mRNAs regulated by *R. prowazekii* sRNAs. Functional categorization of the target genes regulated by sRNAs resulted in identification of genes involved in key pathways of cell division, transport, phagosomal escape, virulence, type IV secretion system, and metabolism. A majority of these pathways are critical for the growth and survival of *Rickettsia* in the host cytoplasm. For example, I identified 33 genes involved in transport mechanisms and potentially regulated by sRNAs, a function important for rickettsial survival *in vivo* as they encode for translocases required for the exchange of ADP with ATP from host cell cytosol (Winkler, 1976; Alexeyev and Winkler, 1999; Gillespie *et al.*, 2007). Following invasion into a host cell, rickettsiae quickly escape into the cytosol by phagosome degradation, and published studies have implicated a role for rickettsial hemolysin C (TlyC) and phospholipase D (Pld) in phagosomal escape (Whitworth *et al.*, 2005; Driskell *et al.*, 2009). I identified two sRNAs, #24 and 27, with the potential to regulate *tlyC* and *pld*, respectively, suggesting an important role for these sRNA in the establishment of infection. A significant number (18%) of predicted target genes were categorized as 'hypothetical proteins', which is not surprising considering that nearly 26% of the 914 *R. prowazekii* genes are still reported as uncharacterized ORFs. As rickettsial genes are further investigated for their functional roles, it is anticipated that most of these

hypothetical proteins will likely be assigned a role in virulence, survival, and pathogenesis during host-pathogen and vector-pathogen interactions.

Table 7. sRNA predictions categorized by nucleotide size.

Predicted sRNA Nucleotide Size							
Rickettsia	30-100	101-200	201-300	301-400	401-500	500-550	Total
Ancestral Group							
R. bellii OSU	33	55	6	5	1	0	100
R. bellii RML	39	43	13	3	2	0	100
R. canadensis	22	14	7	1	1	2	47
Typhus Group							
R. prowazekii Breinl	8	11	5	2	0	0	26
R. prowazekii Madrid E	8	11	5	2	0	0	26
R. typhi	4	5	3	2	1	0	15
Transitional Group							
R. akari	46	43	19	5	1	2	116
R. felis	75	80	25	6	2	0	188
Spotted Fever Group							
R. rickettsii Iowa	53	49	13	8	1	1	125
R. rickettsii Sheila Smith	56	54	17	8	0	0	135
R. africae	62	67	19	10	3	2	163
R. heilongjiangensis	54	69	12	9	3	1	148
R. conorii	62	59	15	4	3	3	146
R. japonica	67	72	22	12	4	1	178
R. massiliae	64	66	19	6	2	0	157
R. peacockii	79	74	17	9	8	4	191

Table 8. sRNA comparison.

	R. bellii OSU	R. bellii RML	R. canadensis	R. prowazekii Madrid	R. prowazekii Breinl	R. typhi	R. felis	R. akari	R. rickettsii SS	R. rickettsii IA	R. conorii	R. africae	R. heilongjiangensis	R. japonica	R. massiliae	R. peacocki
Ancestral																
R. bellii OSU	----------	86	5	4	4	4	58	6	20	24	26	19	23	22	14	32
R. bellii RML		----------	4	4	4	4	60	7	21	24	33	21	28	25	15	37
R. canadensis			----------	3	3	3	11	8	6	5	6	6	5	7	8	3
Typhus																
R. prowazekii Madrid E				----------	31	7	5	4	5	3	5	5	3	2	4	2
R. prowazekii Breinl					----------	7	5	4	5	3	5	5	3	2	4	2
R. typhi						----------	3	3	3	4	3	3	1	1	3	1
Transitional																
R. felis							----------	67	85	65	86	90	78	84	86	91
R. akari								----------	37	34	42	41	41	36	39	38
Spotted Fever																
R. rickettsii SS									----------	120	98	103	88	98	87	98
R. rickettsii IA										----------	103	106	96	107	94	105
R. conorii											----------	126	93	102	94	111
R. africae												----------	100	118	104	121
R. heilongjiangensis													----------	136	82	98
R. japonica														----------	93	109
R. massiliae															----------	92
R. peacocki																----------

Comparison of sRNA predictions and five well-known bacterial sRNAs (6S RNA, α-tmRNA, rpsL_ricks, 4.5S RNA, RNaseP_bact_a) against other rickettsial species and strains. The numbers represent the amount of sRNAs that demonstrate similarity after a BLAST comparison (E-value < $1e^{-5}$).

Table 9. sRNA target predictions.

sRNA Prediction	*Number of Targets*	
	TargetRNA	*CopraRNA*
1*	39	47
2*	3	43
3	4	40
4	11	23
5*	27	50
6	21	49
7	0	53
8	8	37
9*	19	43
10*	32	49
11*	17	49
12	11	44
13	17	49
14	13	37
15	25	50
16	12	45
17	18	51
18	2	54
19	9	45
20	10	39
21*	12	39
22	14	46
23	10	45
24*	24	54
25*	29	36
27	6	37
Total	393	1154

Number of target predictions per each sRNA for *R. prowazekii* strain Breinl. Candidate #26 is not listed, as SIPHT provided no #26 prediction. Asterisks represent those candidates selected for confirmation of expression.

Table 10. sRNA target categorization.

Target Classification	# of Predicted Targets by TargetRNA2	CopraRNA
Cell Division	44	30
Cell Wall	24	83
Metabolism	71	197
Ribosomal Protein	51	129
Virulence	3	17
T4SS	2	28
Other	90	370
Transport	33	67
Phagosome Escape	2	1
Hypothetical Protein	73	232
Total	393	1154

Target genes are classified into ten categories based on either known or hypothetical function for *R. prowazekii* strain Breinl.

Table 11. sRNA predicted promoter locations.

Candidates	Start Position	Stop Position	Strand	-10 box	-35 box
2	163641	163556	Anti-Sense	ATCTAGGAT	TTAATT
5	659164	659057	Anti-Sense	TTGTATTAT	TTTATT
6	644329	644199	Anti-Sense	TAGTAAAAA	TTAGAA
9	457001	456876	Sense	ACTTATCAT	TTGCTG
10	371859	371506	Anti-Sense	TGTTAAAAT	TTTATT
11	308324	308042	Anti-Sense	TTTTGAAAT	TTCTAA
12	306070	305924	Sense	ATGTATATT	TTGATG
21	47692	47542	Sense	GGGTATAAC	ATGACA
22	10482	10278	Anti-Sense	ATGTAAGAT	TTTACT
23	1105018	1104959	Sense	GATCAGAAT	TTCAAA
24	1039473	1039278	Anti-Sense	ATGTAGATT	TTGATT
25	998167	997927	Anti-Sense	GACTAAAAT	TTGCCA

This table outlines those *R. prowazekii* strain Breinl sRNA predictions that had an Mean Expression Value (MEV) ≥1.5. It includes the SIPHT predicted start and stop positions as well as the predicted strand. In addition, it contains the BPROM predicted -10 box and -35 box for the σ^{70} promoters.

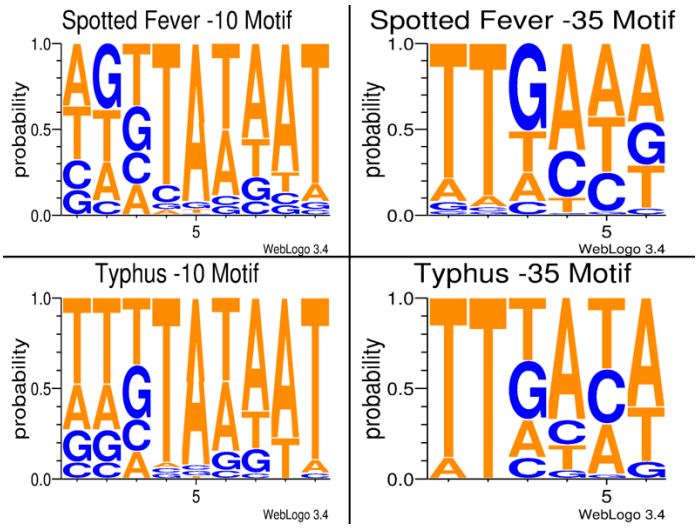


Figure 4. sRNA promoter frequencies.

Conservation diagrams illustrating the probability of a nucleotide in a specific promoter motif position. The left side demonstrates the -10 promoter motif, while the right side is the -35 promoter motif. The upper portion displays the typhus group, while the lower displays the spotted fever group. Both groups have -10 motifs similar to the *E. coli* consensus sequence (TATAAT). On the other hand, the -35 motifs vary when compared to the *E. coli* consensus sequence (TTGACA).

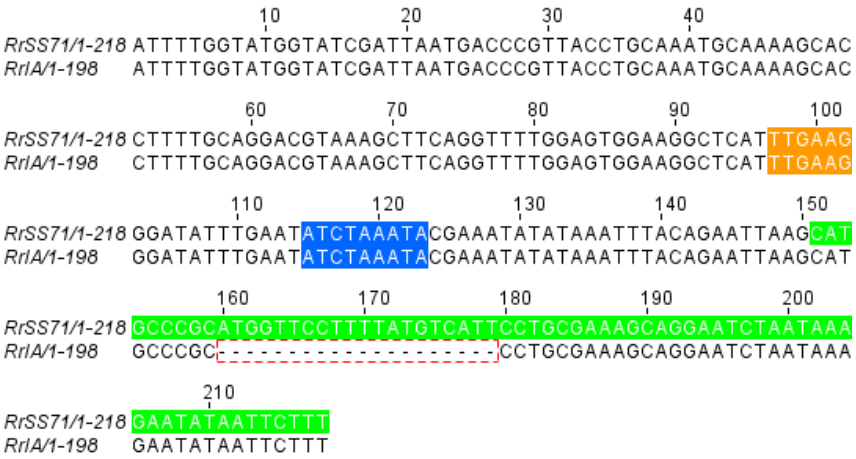


Figure 5. Alignment of *R. rickettsii* strain Sheila Smith sRNA candidate #71.
The sRNA candidate #71, predicted only in R. *rickettsii* strain Sheila Smith but not in strain Iowa and the upstream 150 bp region of predicted sRNA were aligned with the corresponding genomic region from strain Iowa. The predicted -10 box (orange), -35 box (blue), and sRNA sequence (green) are highlighted. A 20 bp deletion observed in the genomic sequence of strain Iowa is shown by the dotted line.

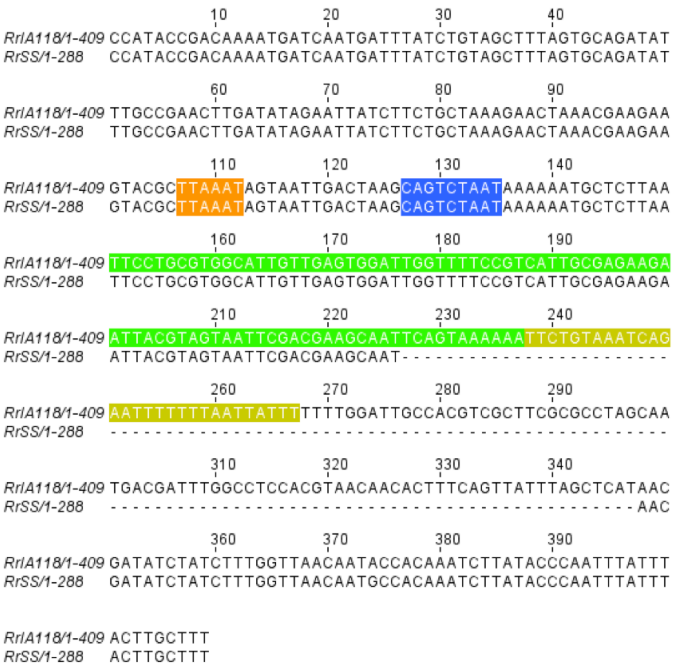


Figure 6. Alignment of *R. rickettsii* strain Iowa sRNA candidate #118.

The sRNA candidate #118, predicted only in R. *rickettsii* strain Iowa but not in strain Sheila Smith and the 150 bp up- and downstream regions of predicted sRNA were aligned with the corresponding genomic region from strain Sheila Smith. The predicted -10 box (orange), -35 box (blue), sRNA sequence (green) and the Rho independent terminator (yellow) are highlighted. A nucleotide sequence absent in the genomic sequence of strain Sheila Smith and mapping to the predicted sRNA and the Rho independent terminator in strain Iowa is shown by the dotted line.

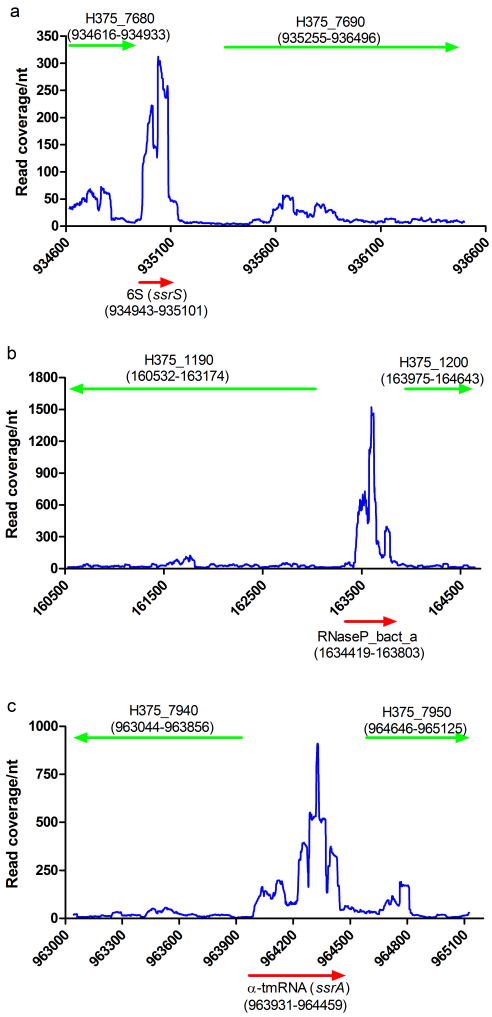


Figure 7. Expression profiles for 6S RNA, RNaseP_bact_a, and α-tmRNA.
Expression profiles for: 6S RNA (Panel a), RNaseP_bact_a (Panel b), and α-tmRNA (Panel c) during *R. prowazekii* infection of HMECs. Y-axis denotes the nucleotide coverage. X-axis shows the genomic location of sRNAs in *R. prowazekii* genome (NC_020993). The sRNAs are shown by red arrows whereas green arrows indicate the flanking up and downstream genes. The numbers in the parenthesis indicate the exact genomic location. The flanking genes and sRNAs are not drawn to scale.

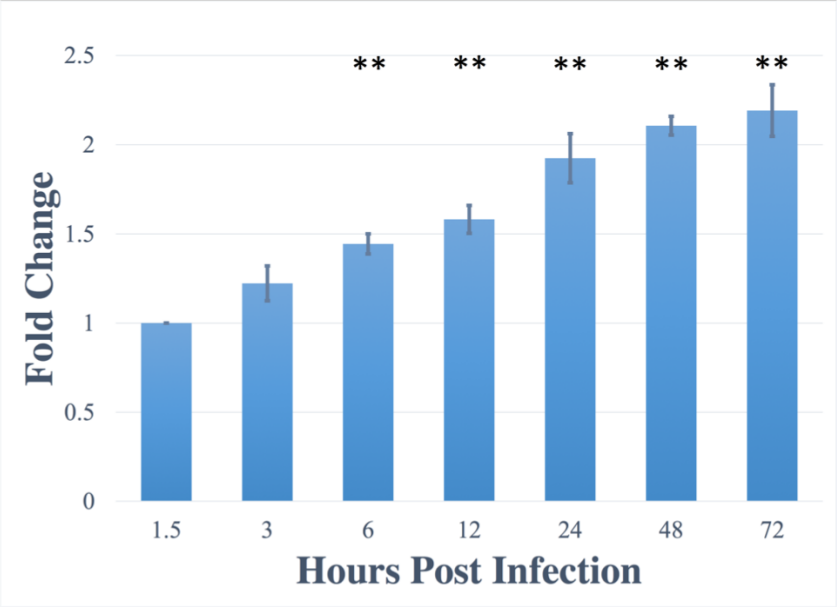


Figure 8. 6S RNA (*ssrS*) expression during host cell infection.
R. prowazekii strain Breinl 6S RNA (*ssrS*) expression was measured during the infection of HMECs over a course of 72 hours (n=5). The expression was normalized to 16S rRNA (endogenous control) and baselined to 1.5 h post infection. Significant increase was observed starting at 6 h post infection. Data are represented as Mean ± SEM. ** $p<0.01$.

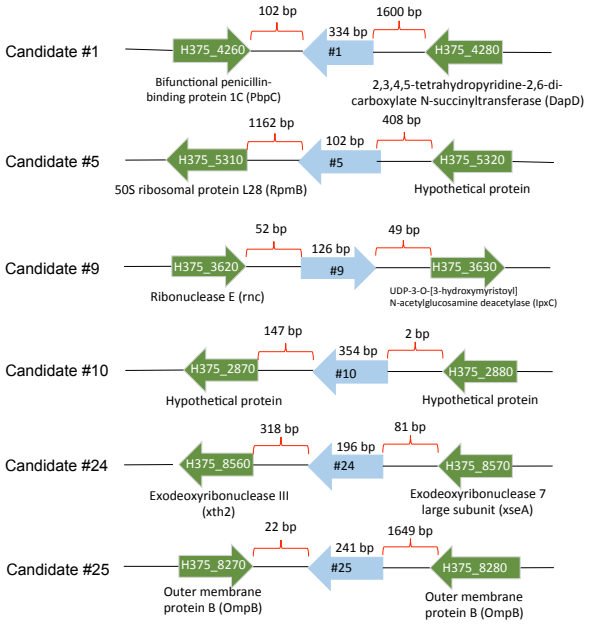


Figure 9. Genomic location of *R. prowazekii* strain Breinl sRNAs.
Schematic representation of sRNAs identified to be expressed in *R. prowazekii* strain Breinl during the infection of HMECs. Green arrows represent the orientation of flanking ORFs in relation to the sRNA depicted by blue arrows. The nucleotide distance between the sRNA and the flanking ORF is shown above the brace.

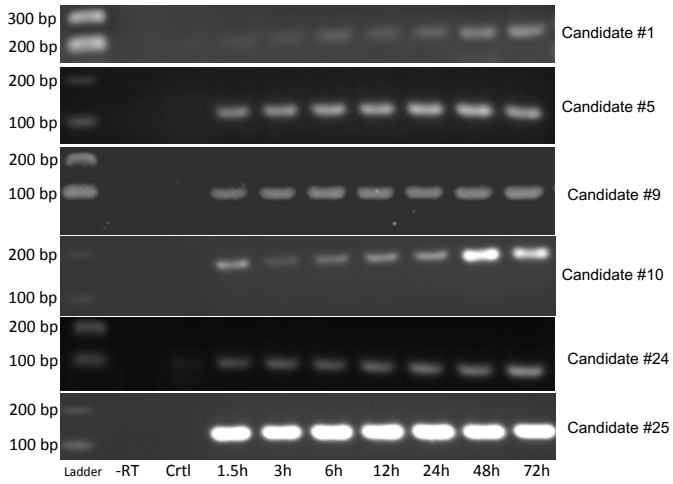


Figure 10. Expression of *R. prowazekii* strain Breinl candidate sRNAs during host cell infection.

R. prowazekii strain Breinl sRNA candidates #1, #5, #9, #10, #24, and #25 were tested for their expression during infection of HMEC by RT-PCR (n=3). The band sizes shown on the left side correspond to the 100 bp DNA ladder (New England Biolabs). The lane 2 (-RT) is a "no reverse transcriptase" control, while the lane 3 (Crtl) is an uninfected HMEC control. Lanes 4 through 10 are the samples from *R. prowazekii* strain Breinl infected HMECs from 1.5h to 72h post infection. All the tested sRNA candidates showed expression during host cell infection.

Chapter 4 Identification and Validation of Novel Candidate sRNAs within *R. prowazekii* strain Breinl[3]

INTRODUCTION

Sequencing of a pathogen's genome is an effective umbrella approach to identify unknown genotypic/phenotypic traits and to establish a platform for dissecting and deciphering gene function. Ready availability of a number of rickettsial genomes has had a dramatic impact on our understanding of their genetic diversity, genomic architecture, gene identification and function, and mechanisms of pathogenesis. As the first rickettsial genome to be sequenced and published (Andersson *et al.*, 1998), *R. prowazekii* was found to carry a rather high amount (approximately 24%) of non-coding DNA (Andersson *et al.*, 1998; Holste *et al.*, 2000) and an AT-rich genome with a GC content of 29.1%, suggesting genomic reduction and gene neutralization due to obligate intracellular parasitism. While intergenic regions in other bacteria harbor small non-coding RNAs, their presence in *Rickettsia* species remained an open question until recently, when I predicted a number of sRNAs in rickettsial genomes using complementary computational approaches (Schroeder *et al.*, 2015). Using infection of human microvascular endothelial cells (HMECs) with *R. prowazekii* as an experimental model system, I further determined the presence of six novel *trans*-acting sRNAs, but a limitation of this study was the exclusion of potential *cis*-acting sRNAs and possibly other novel *trans*-acting sRNAs. Here, I report on the identification, validation, and characterization of both *cis*-acting and

[3] The work mentioned in this chapter is based on the work published: Schroeder, C. L., H. P. Narra, A. Sahni, M. Rojas, J. Patel, K. Khanipov, et al. (2016). Identification and characterization of novel small RNAs in *Rickettsia prowazekii*. *Frontiers in Microbiology*, 7:859. The Creative Commons License can be accessed at http://creativecommons.org/licenses/by/4.0/

additional novel *trans*-acting sRNAs in *R. prowazekii*. I identified 35 novel trans-acting and 23 cis-acting sRNAs through next generation sequencing and confirmed the expression of four novel sRNAs in addition to well-known noncoding sRNAs, namely, α-tmRNA, RNaseP_bact_a, *ffs*, and 6S RNA. I further analyzed the validated sRNAs using experimental and bioinformatics techniques to determine their transcriptional start sites, upstream promoter motifs, and potential target genes.

RESULTS

Identification of R. prowazekii sRNAs through RNA Sequencing.

In a recently published study, I predicted the existence of 26 candidate sRNAs within the *R. prowazekii* strain Breinl genome using a web-based interface SIPHT. Upon further analysis, 12 of these candidates were found to have an MEV of >1.5 and six were confirmed through RT-PCR (Additional File 3) (Schroeder *et al.*, 2015). A limitation of this study, however, was that programs such as SIPHT only survey the intergenic regions and do not screen regions antisense to ORFs, thus failing to identify *cis*-acting sRNAs. To investigate the presence of *cis*-acting sRNAs and to perform a deeper analysis of the transcriptome to identify additional *trans*-acting sRNAs that do not meet the parameters of the SIPHT program, I performed next generation sequencing to identify and catalogue all sRNAs in *R. prowazekii*.

RNA sequencing of HMECs infected with *R. prowazekii* for 3 h (to enable entry and infection) resulted in approximately 42 to 46 million total reads, whereas 27 to 29 million total reads were obtained for RNA isolated from cells infected for 24 h. Of these, approximately 930,000 to 2 million reads mapped to the *R. prowazekii* genome at 3 h and

a total of 1 to 4 million reads corresponded to *R. prowazekii* at 24h. This is in agreement with a recent demonstration that intracellular organisms, such as *Rickettsia* species, constitute only 5% of extracted total RNA, whereas the remaining 95% belongs to eukaryotic host cells. Further analysis revealed that about 95% of bacterial RNA was comprised of rRNAs and tRNAs and only the remaining 5% included mRNAs and sRNAs. This, in essence, translates to a ratio of ~1:400 bacterial mRNAs and sRNAs in the preparations of total cellular RNA from infected host cells (Westermann *et al.*, 2012). Despite efficient removal of eukaryotic polyadenylated transcripts and ribosomal RNAs through microbial enrichment protocols, the process is limited in removing tRNAs, eukaryotic noncoding RNAs, and mitochondrial RNAs, resulting in possible interference. Consequently, only 2-5% of extracted total RNA generally maps to intracellular bacterial genomes (Westermann *et al.*, 2012).

RNA-seq based search of the transcriptome identified a total of 70 candidate *trans*-acting (intergenic) and *cis*-acting (antisense) sRNAs that were either expressed at 3 and/or 24h post-infection (Additional File 3). All sequences are available in the Bacterial Small RNA Database (Li *et al.*, 2013) and GenBank (accession number KX215777 through KX215846). A representative selection of RNA read coverage plots for *trans*- and *cis*-acting sRNAs is presented in Figure 1 and Figure 2, respectively. Amongst the newly identified candidates, 35 candidates were *trans*-acting and another 23 were *cis*-acting sRNAs. The sizes of the candidates identified ranged from 59 bp to 585 bp, with an average of 233bp. Interestingly, four *trans*-acting candidates (named *Rickettsia prowazekii* small RNAs [*Rp*_sR]5, *Rp*_sR45, *Rp*_sR50, and *Rp*_sR66) were found to have sequence homologs in other rickettsial species. All four had strong (~90%)

homologs in the typhus group *R. typhi,* while candidate *Rp*_sR5 also shared a strong homolog in *R. felis*, a transitional group species. Otherwise, all remaining *trans*-acting sRNAs were unique to *R. prowazekii*. With the exception of two candidates (*Rp*_sR3 and *Rp*_sR18), *cis*-acting sRNA candidates generally shared considerable homology to their sRNA counterparts in rickettsial species outside of the typhus group. The corresponding ORF for *Rp*_sR3 is H375_160 (a transcription-repair coupling factor), which is present in all other sequenced *R. prowazekii* strains. For *Rp*_sR18, the corresponding ORF is a single-stranded DNA-specific exonuclease (H375_870), which is also conserved in other *R. prowazekii* and *R. typhi* strains. Considering that the intergenic regions are more likely to undergo dynamic changes in their sequences over time and *cis*-acting sequences tend to remain conserved by virtue of their location on the anti-sense strand of an ORF, the high number of candidates sharing homology outside of the typhus group is not surprising. Accordingly, 17 *cis*-acting sRNAs were present on the ORFs conserved in all sequenced rickettsial species, and 4 sRNAs were encoded from an antisense strand representing hypothetical and ORFan proteins (ORFs with no known homologs within current databases). Interestingly, candidate *Rp*_sR12 had no homologs in the spotted fever group, but it did share a homolog in the ancestral group, typhus group, and transitional group. Further, candidate *Rp*_sR28 had homologs in the typhus group and transitional group, but no homologs were found in the ancestral group or the spotted fever group. Candidate *Rp*_sR22 shared homologs in the typhus group and the spotted fever group, but not the transitional group. It is important to note that homology was not necessarily observed for all rickettsial species of a particular group. For example,

homologs of candidate *Rp*_sR22 were found in the spotted fever group species *R. montanensis* and *R. japonica*, but not *R. rickettsii* or *R. conorii.*

Experimental Validation of Candidate sRNAs

To verify expression of candidate sRNAs, Northern blots were carried out with strand-specific [α-^{32}P] UTP-labeled RNA oligonucleotide probes specific to the corresponding candidate with enriched rickettsial RNA from *R. prowazekii*-infected HMECs at 3 and 24h post-infection. Two *trans*-acting sRNAs (*Rp*_sR17 and *Rp*_sR60) and two *cis*-acting sRNAs (*Rp*_sR34 and *Rp*_sR47) identified from the RNA-seq data were selected on the basis of their genomic location, number of RNA reads, strand orientation, as well as the orientation of corresponding upstream and downstream ORFs. Candidate *Rp*_sR67 [identified earlier with the SIPHT program as Candidate #25 (Schroeder *et al.*, 2015)] was also selected for verification via Northern blot analysis. In addition, four other highly conserved bacterial sRNAs were chosen. These included 6S RNA (*ssrS*), 4.5S RNA (*ffs*), RNaseP_bact_a, and α-tmRNA (*ssrA*). 6S RNA was only probed at 24 h post-infection as my previous data demonstrated that rickettsial 6S RNA was most abundantly expressed at this time-point (Schroeder *et al.*, 2015). Strong and distinct RNA bands were observed for six probed sRNAs except for α-tmRNA, which had a light but distinct band (Figure 13). *Rp*_sR34 and *Rp*_sR60 could not be detected by Northern blot analysis. Interestingly, the blots for RNaseP_bact_a, *Rp*_sR17, and *Rp*_sR47 revealed multiple bands, suggesting processed sRNA transcripts. As these bands are relatively similar in size, the fact that I did not see these differences in the RNA-seq data is not surprising since the libraries were generated from both primary and

processed transcripts. Because the RNA probes utilized in my experiments were strand-specific and the detected signals were at their predicted sizes, the Northern blot analysis validates and confirms independent transcription of the candidate sRNAs during host cell infection.

Since *Rp*_sR60 was not detected using Northern blot, RT-PCR was done as an alternative approach to confirm expression. Total RNA from *R. prowazekii*-infected HMECs clearly demonstrated expression of *Rp*_sR60 at both 3h and 24h post-infection. Figure 14 shows a representative agarose gel with 16S RNA serving as an endogenous control. A similar approach was not amenable for the detection of *Rp*_sR34 because of its *cis*-acting nature and the possibility of confounding amplification of its corresponding ORF transcript. These results, nevertheless, provide clear evidence for the expression of *Rp*_sR60 during *R. prowazekii* infection of HMECs.

Nucleotide Sequencing of Validated sRNAs by RLM-RACE

Northern blot analysis only provides information regarding the levels of expression and the transcript size, but is unable to reveal the exact transcription start site. Therefore, to determine the transcription start sites (TSS) of confirmed novel sRNAs (*Rp*_sR17, *Rp*_sR34, *Rp*_sR47, *Rp*_sR60, and *Rp*_sR67) and 6S RNA, 5' RLM-RACE was employed. RLM-RACE provides a distinct advantage over the traditional 5' RACE method because it only amplifies primary transcripts, and not processed RNAs, for subsequent sequencing, which then permits the determination of an exact transcription start site. The TSSs for the novel *R. prowazekii* sRNAs, confirming the potential for their independent expression, are listed in Table 12. Furthermore, the TSSs deciphered from

the RLM-RACE approach were found to correspond to transcriptomic data from RNA-seq experiments and to closely align with the initiation loci of sRNAs' expression. This correlation, thus, further supports independent expression of the *Rp*_sRs during infection of target host cells.

Molecular Analysis of 6S RNA and Novel R. prowazekii sRNAs

In order to analyze each confirmed sRNA for potential promoter motifs (-10 and -35 sites), 150 nucleotides upstream of the identified transcription start site were submitted to the web-based program, BPROM. This range was chosen keeping in mind that approximately 80% of known σ^{70} promoters in *Escherichia coli* are located within 150 nucleotides of the transcription start site (Huerta and Collado-Vides, 2003). Upon predicting σ^{70} promoters for all novel sRNAs and 6S RNA, a deeper analysis demonstrated that the -10 and -35 motifs were an average of 60 nucleotides and 82 nucleotides upstream of the transcription start site, respectively. This distance ranged from 36 to 93 bp upstream for the -10 motif and 56 to 114 bp upstream for the -35 motif (Table 12). For *E. coli,* the optimum distance between the -10 and -35 motifs has been reported to be 17±1 nucleotides (Harley and Reynolds, 1987; Mitchell *et al.*, 2003), but my data indicate that the average distance for *R. prowazekii* is 21 nucleotides. This result is in close agreement with my previous findings for sRNAs predicted by SIPHT in spotted fever and typhus group rickettsiae (Schroeder *et al.*, 2015). The predictions, which have a consensus of TATAAT, are identical to accepted -10 consensus sequence for *E. coli* (Harley and Reynolds, 1987). The first and second position of the -10 motif is conserved with T and A at 100% and 85%, respectively. The third and fourth positions

are approximately 50% T and A. The fifth position is 85% A, while the sixth position is 100% T. On the other hand, the rickettsial -35 motif, predicted as TTGCAA, has notable differences compared to the -35 consensus sequence for *E. coli*, reported as TTGACA (Harley and Reynolds, 1987). For the rickettsial sRNAs, the first three positions of the -35 motif are TTG nearly 82% to 100% of the time. The fourth position is a C with approximately 50% probability as opposed to an A in *E. coli* consensus sequence. The fifth position for *E. coli* is C, but it is A or T in *R. prowazekii*. The final position was an A approximately 70% of the time (Figure 15).

Candidate sRNA Target Identification

Two independent programs, TargetRNA2 and IntaRNA, were used to further understand the regulatory roles of the confirmed novel *trans*-acting sRNAs (*Rp*_sR17, *Rp*_sR60, and *Rp*_sR67). These programs identify sRNA:mRNA interactions by assessing base pairing potential based on a Smith-Waterman dynamic and by assessing interaction sites and seed regions, respectively (Busch *et al.*, 2008; Tjaden, 2008). By using TargetRNA2, a total of 72 protein-coding targets were identified to be potentially regulated by sRNAs. Conversely, IntaRNA predicted a total of 122 targets with a $p \leq 0.05$. A detailed analysis revealed common target predictions by both programs (Table 13). For *Rp*_sR60 and *Rp*_sR67, a total of 7 and 6 common protein targets were found, respectively. Three common targets predicted for *Rp*_sR67 have functional roles in tRNA synthesis and another one participates in ATP synthesis, while the other common targets were hypothetical or uncharacterized proteins. Interestingly, *Rp*_sR60 has two protein-coding mRNA targets involved in ribosomal protein synthesis and three targets

involved in cell metabolism. In contrast, *Rp*_sR17 has a single common predicted target, namely protein translocase subunit, *secD*. This protein facilitates secretion across the inner membrane of Gram-negative bacteria. The target genes regulated by each sRNA are listed in Additional File 4. Targets for the *cis*-acting sRNAs were not predicted as previous data demonstrate that *cis*-acting sRNAs tend to interact with their associated ORFs. This suggests the potential for *Rp*_sR34 interaction with H375_2470, a hypothetical protein, whereas *Rp*_sR47 may base-pair and regulate the expression of H375_3890, an ORF encoding for a carboxyl-terminal protease on the complementary strand. Although a large percentage of *R. prowazekii* genes either code for hypothetical proteins or remain uncharacterized, a majority of predicted mRNA targets for novel sRNAs have an assigned function. It is reasonable to posit, therefore, that sRNAs may have a regulatory influence on at least some of the predicted targets and are likely to be functional in *R. prowazekii*.

DISCUSSION

Despite recent developments and tremendous progress in the identification of bacterial sRNAs and discoveries uncovering their novel roles as important post-transcriptional regulators, rickettsial sRNAs remain largely unknown and uncharacterized. A recent study from my laboratory reported on the predictive analysis of rickettsial sRNAs by SIPHT and an initial confirmatory analysis of their presence and expression in infected host cells by RT-PCR (Schroeder *et al.*, 2015). Here, I report on a transcriptome-wide analysis focused on the identification of entire repertoire of both *trans*-acting and *cis*-acting sRNAs in the virulent Breinl strain of *R. prowazekii* expressed

during the infection of host cells *in vitro*. I identified a total of 58 novel sRNAs through RNA-seq and confirmed the expression of four sRNAs along with 6S RNA, α-tmRNA, RNaseP_bact_a, and 4.5S RNA. I further performed RLM-RACE to identify their transcription start sites and computational promoter analysis on the confirmed sRNAs. Also, target predictions using two independent programs, TargetRNA2 and IntaRNA, have revealed probable mRNA targets for each *trans*-acting sRNA.

Most sRNAs described to date have been found in model organisms, such as *E. coli*, *Salmonella enterica* serovar Typhimurium, and *Staphylococcus aureus* (Sharma and Heidrich, 2012). These sRNAs have been found to control gene expression in bacterial virulence, stress response, and other necessary functions and thus play a critical role in bestowing the adaptation skills that allow the organism to launch a quick response to environmental changes (DiChiara *et al.*, 2010; Sharma and Heidrich, 2012). For example, *S. aureus* encodes SprD, a regulatory RNA that targets the Sbi immune-evasion molecule. Upon interaction with the 5' region of the *sbi* mRNA, the central region of the sRNA interacts with the target translational start site and negatively regulates expression of the Sbi immune-evasion molecule, which impairs the host's innate and adaptive immune response (Chabelskaya *et al.*, 2010). Similarly, biofilm formation in *E. coli* is increased with an overexpression of McaS (multi-cellular adhesive sRNA) and compromised in strains lacking the *mcaS* gene. The McaS is responsible for repressing the curli biogenesis gene (extracellular proteinaceous attachment proteins) (*csgD*), activating the flagella synthesis regulator (*flhD*) to increase motility, and decreasing export of the polysaccharide through porin expression (*pgaA*) (Thomason *et al.*, 2012). These sRNAs have critical implications on bacterial virulence and survival. Although

not yet characterized in as much detail, sRNAs are now beginning to be described in a wide range of *α-proteobacteria*, including *Coxiella*, *Bartonella*, and *Rickettsia* species (Warrier *et al.*, 2014; Schroeder *et al.*, 2015; Tu, 2015).

Previous studies demonstrate that *R. prowazekii* relies on leaky transcriptional termination for host adaptation and survival (Woodard and Wood, 2011). Therefore, to ascertain whether the candidates represent independent transcripts or are simply the result of leaky termination, the MEV for each of the sRNA candidates that were in the same orientation as the closest upstream ORF was calculated and compared to the MEVs calculated for the respective 50-base flanking regions. Of the 58 candidate sRNAs, 22 candidates were identified to be in the same orientation as the closest upstream ORF. Each candidate had an MEV of ≥5 when compared to the surrounding 50-base flanking regions with the only exception of candidate *Rp*_sR17, which had an MEV of ≥2 when compared to its flanking region. Even though it did not surpass the cut-off threshold of 5-fold used in this study, expression of sRNA candidate *Rp*_sR17 is quite likely to be independent from its neighboring ORF's and the low MEV may be due to its lower level of expression during infection of HMECs. The remaining 33 candidate sRNAs were in an orientation opposite their closest upstream ORF. For this situation, it is reasonable to expect that evidence for sRNA expression is not an outcome of leaky termination, as the reads corresponding to candidate sRNAs map onto the opposite strand. These data support the notion that identified candidate sRNAs are not the consequence of leaky termination, but instead represent bona fide transcripts expressed independently of their flanking ORFs.

In this study, I have identified a total of 32 novel sRNAs to be specifically present in the *R. prowazekii* genome, while only 26 sRNAs were shared among species belonging to other rickettsial groups (Additional File 3). The occurrence of a vast number of *R. prowazekii*-specific sRNAs is not surprising, considering that rickettsial genomes have undergone genome degradation and rearrangements resulting from horizontal gene transfer (HGT), transposon mutagenesis, and pseudogenization (Andersson *et al.*, 1998; Fournier *et al.*, 2009; Gillespie *et al.*, 2012). For instance, the genome of the *Rickettsia* endosymbiont of *Ixodes scapularis* (REIS) is replete with mobile genetic elements and conjugative plasmids resulting in acquisition of nearly 32% of the ORFan genes unique to REIS. It is presumed that large tracks of unique genes along with their IGRs may have been acquired from other unknown organisms with AT-rich genomes similar to rickettsiae (Gillespie *et al.*, 2012). Species-specific sRNAs arising from genome rearrangements, deletions, and point mutations have been reported in several bacteria. For instance, *E. coli* and *S. enterica* serovar Typhimurium, despite having evolutionarily conserved genomes, are known to encode species-specific EcsR1 and SesR2 sRNAs, respectively (Raghavan *et al.*, 2015). SesR2 has been shown to transcribe from an IGR resulting from a genome rearrangement due to phage translocation. Furthermore, point mutations resulting in the evolution of new RpoD promoters are also presumed to transcribe novel transcripts resulting in the biogenesis of species-specific non-coding RNAs (Mendoza-Vargas *et al.*, 2009).

Although next generation sequencing yields convincing data suggesting expression of selected *R. prowazekii* sRNAs during infection of HMECs, an important next step was to validate their expression via independent experimental methods

considered to be the 'gold standard' in the fields of RNA biology in general and bacterial sRNAs in particular. Accordingly, I chose four novel sRNA candidates for validation through Northern blot analysis with strand-specific RNA probes and 5' RLM-RACE. Previously, qRT-PCR based analysis of *R. prowazekii*-infected HMECs demonstrated that 6S RNA expression was significantly higher at 24 h post-infection in comparison to the same at 1.5 h (Schroeder *et al.*, 2015). In the present study, Northern blotting confirmed abundant expression of 6S RNA at 24 h post-infection as expected. Importantly, Northern blots and RT-PCR further revealed clear expression of three selected novel sRNA candidates (*Rp*_sR17, *Rp*_sR47, and *Rp*_sR60) at 3 and/or 24 h post-infection. In regards to the well-known sRNAs (RNaseP_bact_a, α-tmRNA, 4.5S RNA, and 6S RNA), there was clear expression at 3 h and 24 h post-infection. The successful sequencing and determination of transcriptional start sites for each novel sRNA further demonstrates that *R. prowazekii* does, indeed, express these sRNAs during host cell infection. When the RLM-RACE data are overlaid with the RNA-seq data, the transcription start sites, as determined by RLM-RACE, correspond to the start of the sequencing reads with the exception of *Rp*_sR60, in which case the transcription start site determined by RLM-RACE mapped towards the middle of the observed RNA-seq reads. This could potentially be an outcome of: (i). occurrence of two novel small RNAs that may either have a significant overlap or may be located within extremely close proximity of each other in the *R. prowazekii* genome, (ii). presence of an unannotated ORF upstream of the sRNA TSS, or (iii). *Rp*_sR60 could be a riboswitch containing sRNA. Translation of the sequence upstream of the TSS identified by RACE revealed the existence of a predicted hypothetical ORF between positions 844285 and 844416.

However, the -10 and -35 motifs were absent upstream of this predicted ORF. Riboswitch sRNAs are present in the 5' untranslated region of the coding genes and regulate expression of the downstream gene in response to metabolic stimuli. Several riboswitches have been well characterized in different bacterial species, and *Bartonella,* a facultative intracellular pathogen, is shown to express at least 9 riboswitch sRNAs (Brt 1-9) upstream of different helix-turn-helix XRE genes (Tu, 2015). Most recently, EutX and Rli55 sRNAs belonging to *Enterococcus faecalis* and *Listeria monocytogenes*, respectively, were shown to contain riboswitches regulated by ethanolamine. The EutX is expressed as two independent transcripts depending on the presence or absence of ethanolamine and cofactor adenosylcobalamin (AdoCbl). The larger transcript (~300 bases) is predominantly expressed in the presence of ethanolamine, but not AdoCbl. In contrast, an ~150 nucleotide transcript is present only in the presence of AdoCbl, indicating metabolite-dependent variations in the sRNA transcript length (DebRoy *et al.*, 2014; Mellin *et al.*, 2014). Further studies on the *Rp*_sR60 genomic region are, therefore, necessary to resolve the differences in my RNAseq and RLM-RACE based annotation of TSS. Taken together, the results from Northern blot analysis and RLM-RACE sequencing as two independent approaches not only substantiate the presence and expression of the sRNA candidates under investigation in *R. prowazekii*, but also allow for the determination of their transcriptional start sites within the genome.

Using the web-based program RNAfold (Hofacker, 2002), a prediction of the secondary structure for *R. prowazekii* 6S RNA based on the RLM-RACE data and the RNA-seq data reveals that it is quite similar to other published bacterial 6S RNAs (Figure 16) (Barrick *et al.*, 2005). The structure is composed of a single central strand, the ends

of which contain either a closed stem or a terminal loop. The purpose of the central bubble is to mimic an open promoter complex on a DNA template (Wassarman and Storz, 2000; Wassarman, 2007). This bubble is responsible for RNA polymerase binding to the 6S RNA molecule and eventually becoming sequestered. Based on its pattern of expression and the predicted secondary structure, it is quite likely that 6S RNA in *R. prowazekii* is functional. I also predicted the secondary structures of novel rickettsial sRNAs using the determined start site and a predicted termination location in Additional File 5. As an example, *Rp*_sR67 predictably contains three arms organized around a central bulge. The sRNA secondary structures are predominantly projected as indicators of evolutionary conservation/relationships and structure-function prediction based on the thermodynamics of stable structures (Mathews *et al.*, 2010). Distinct mechanisms of regulation are employed by *Vibrio harveyi* quorum sensing Qrr sRNA depending on the base pairing of the nucleotides in different stem loops that are involved in interactions with the target gene. The binding of *luxM* and *aphA* target genes to the first stem loop of Qrr sRNA results in the degradation of Qrr sRNA. The strong binding of luxO to the second stem loop, on the other hand, leads to the sequestration of the sRNA and catalytic repression of Qrr occurs following relatively weak interactions between the sRNA and luxR mRNA (Feng *et al.*, 2015). The prediction of multiple target genes regulated by the *R. prowazekii* sRNA repertoire (Table 13). coupled with the existence of complex RNA fold structures such as several predicted stem loops suggests the potential for their involvement in diverse regulatory mechanisms. Hence, the details of their functions and the mechanisms of action are critically important topics indeed for further detailed investigations.

The discovery of small RNAs in a wide range of bacterial pathogens and the obvious need for further definition of their functions served as a stimulus for the development of algorithms to predict their possible targets. As such, a variety of computational programs have been developed for the prediction of small RNA targets. For this study, I chose TargetRNA2 and IntaRNA to predict the potential mRNA targets for *trans*-acting sRNAs in *R. prowazekii*. An important consideration for *trans*-acting sRNAs originating from intergenic regions is that their targets can be located elsewhere in the genome and are not necessarily restricted to adjoining or neighboring ORFs. As expected, the predicted targets for *Rp*_sR17, *Rp*_sR60, and *Rp*_sR67 primarily include bacterial metabolism and other housekeeping functions. However, a confounding factor is that a number of potential targets are either hypothetical or uncharacterized genes, rendering the possible downstream functional implications of these sRNAs difficult to analyze and interpret. It must also be noted that most *trans*-acting small RNAs require the use of a chaperone molecule to facilitate sRNA:target binding due to a limited nucleotide similarity. The most prominent among the sRNA chaperones is Hfq; a small hexameric RNA-binding protein, whose role as a post-transcriptional regulator has been recognized relatively recently (Chao and Vogel, 2010). Hfq-dependent sRNAs usually repress translation and/or increase target destruction through ribonuclease E (RNase E) activity (Chao and Vogel, 2010), although there are some examples in which Hfq-dependent sRNAs stabilize their target mRNAs. Interestingly, both typhus group (*R. prowazekii*) and spotted fever group (e.g., *R. rickettsii, R. conorii*) are not known to encode for a canonical *hfq*, nor has another chaperone molecule been described in *Rickettsia* species (Sun *et al.*, 2002; Chao and Vogel, 2010). Nearly all members of

Enterobacteriaceae encode the *hfq* gene, but nearly 50% of other bacteria are know not to code for the gene (Sharma and Heidrich, 2012). These include *Streptococcus, Mycobacteria, Helicobacter,* and *Chlamydia*, which all have validated small, noncoding RNAs. This suggests the use of an alternative chaperone molecule to facilitate rickettsial *trans*-acting sRNA interactions with target mRNAs. Indeed, potential non-Hfq chaperone molecules have been described in *Helicobacter pylori* and *Mycobacterium tuberculosis*.

Because *cis*-acting sRNAs exhibit perfect nucleotide similarity, the expected target would be the corresponding ORF on the opposite strand. In this context, H375_2470 as the corresponding ORF for *Rp*_sR34 has been annotated as a hypothetical protein, and thus it is difficult indeed to hypothesize the significance or role of this sRNA for *R. prowazekii* until the function of this protein in the rickettsial lifecycle has been ascertained. A BLAST search reveals that H375_2470 is conserved in rickettsial genomes; therefore, it is likely a functional protein. On the other hand, *Rp*_sR47 is antisense to a known carboxyl-terminal protease (H375_3890). Carboxyl-terminal proteases (CTPs) are a group of serine proteases, including tail-specific proteases (TSPs), that have been recognized as critical players in bacterial protein processing (Hoge *et al.*, 2011; Seo and Darwin, 2013). The functions of CTPs include post-translational modification, maturation, and/or disassembly/degradation of proteins performing basal physiological functions and virulence factors (Lad *et al.*, 2007; Hoge *et al.*, 2011). *Chlamydia trachomatis,* an obligate intracellular human pathogen, has recently been described to secrete a tail-specific protease CT441, which degrades p65 and disrupts the NF-κB pathway of host antimicrobial and inflammatory responses (Lad *et al.*, 2007).

Similarly, *R. rickettsii* interacts with inactive NF-κB of the endothelial cell cytoplasm in a 'cell-free' system resulting in its activation as evidenced by increased DNA-protein binding in a mobility shift assay (Sahni *et al.*, 1998; Sahni *et al.*, 2003). Furthermore, *Pseudomonas aeruginosa*, an opportunistic human pathogen, relies on CtpA for normal function of its type 3 secretion system (T3SS). The T3SS is essential for cytotoxicity towards host cells and is a vital virulence factor in mouse models of acute pneumonia. Conversely, up-regulation of CtpA in *P. aeruginosa* induces the expression of an extracytoplasmic function sigma factor regulon, resulting in an attenuated phenotype in rat models of chronic lung infection (Seo and Darwin, 2013). In the facultative intracellular pathogen *Brucella suis, ctpA*-deficient strains display altered morphology, increased cell size, and partial dissociation of cell membrane from the envelope and are cleared nine weeks post-inoculation in a BALB/c mouse model, suggesting that CtpA is critical for survival in macrophages (Bandara *et al.*, 2005). Since my RNA-seq data indicate simultaneous expression of both *Rp*_sR47 and H375_3890, I hypothesize that *Rp*_sR47 interaction with H375_3890 may lead to the stabilization of the transcript for efficient translation. Further evaluation of this interaction and its functional implications in *R. prowazekii* are currently ongoing.

The field of bacterial pathogenesis is rapidly evolving and expanding with enhanced appreciation of the vastness of virulence factors and effector sRNAs. Here, I report on the identification of a number of sRNAs encoded and expressed by *R. prowazekii* during *in vitro* infection of vascular endothelial cells, the primary target cell-type during human disease. I have further substantiated the potential implications of this new regulatory paradigm in rickettsial biology via determination of transcription start

sites of select novel rickettsial sRNAs and predictive analysis of their target genes involved in the pathways of bacterial maintenance, growth, and survival, and mechanisms of pathogenesis. In summary, this study provides the very first glimpse of the non-coding transcriptional landscape of *R. prowazekii* in target host cells.

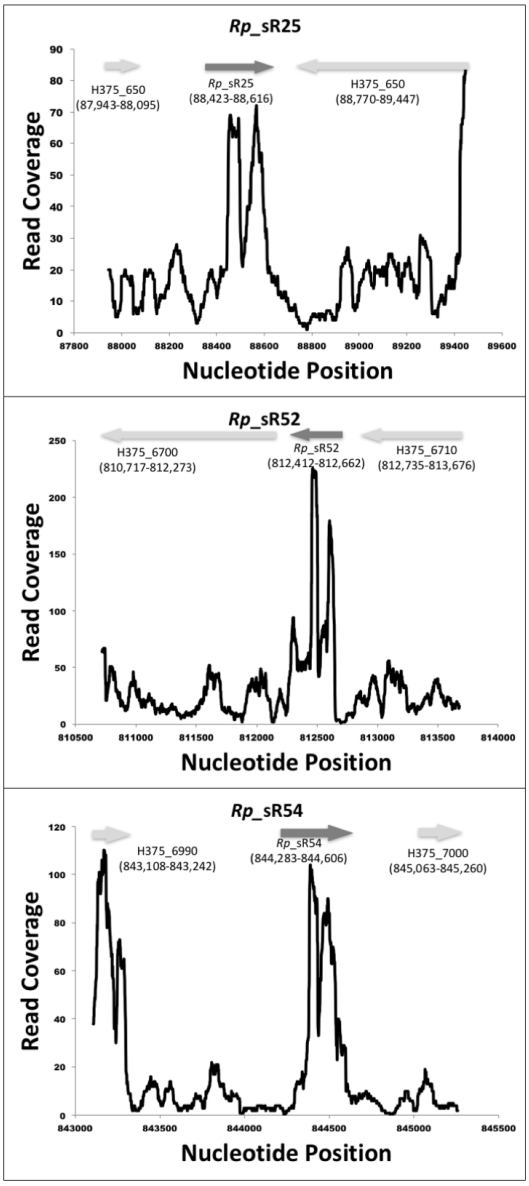


Figure 11. Identified novel *trans*-acting candidate sRNAs.
Shown are the coverage plots for selected *trans*-acting sRNAs. Nucleotide positions within the genome are indicated on X-axis and the Y-axis displays the number of reads for that particular nucleotide position. The dark grey arrow represents the sRNA. The light grey arrows represent the orientation of upstream and downstream ORFs, respectively.

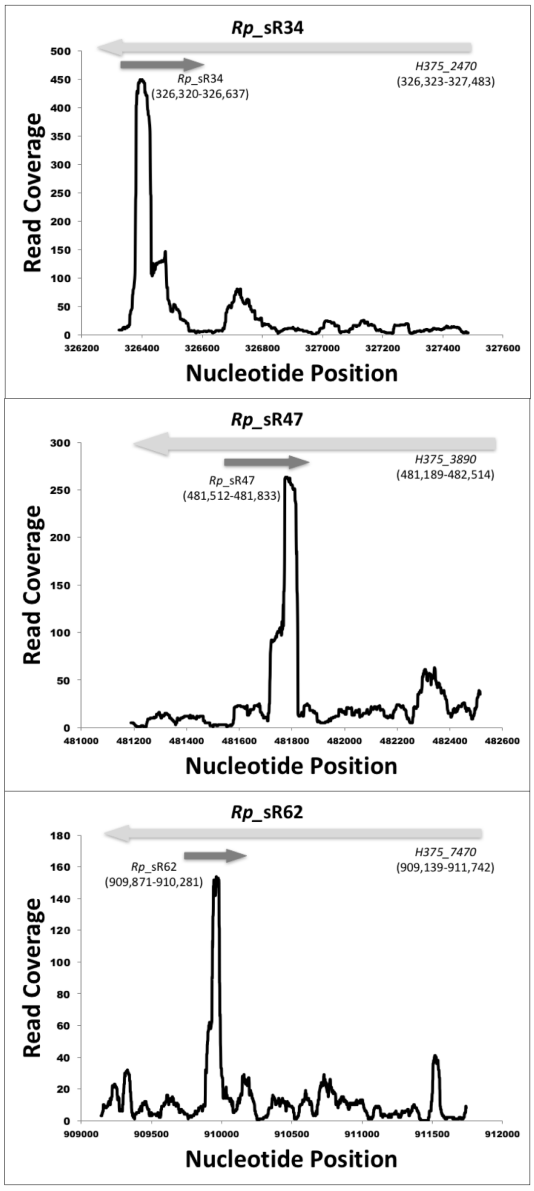


Figure 12. Identified novel cis-acting candidate sRNAs.

Shown are the coverage plots for selected *cis*-acting sRNAs. Nucleotide positions within the genome are indicated on X-axis and the Y-axis displays the number of reads for that particular nucleotide position. The dark grey arrow represents the sRNA. The light grey arrows represent the orientation of the respective ORF.

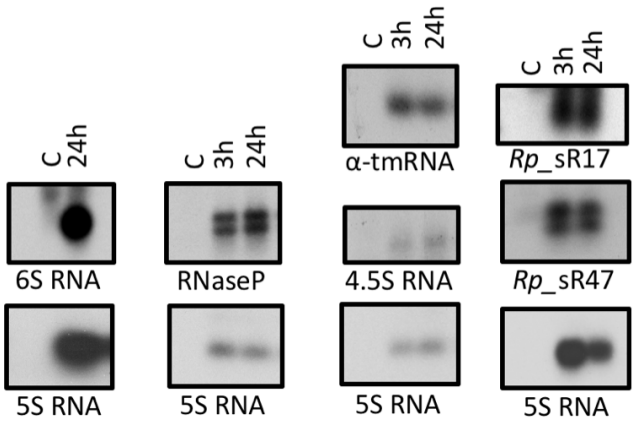


Figure 13. Northern blot analysis.

A representative image of the Northern blots for candidate sRNAs is shown. Northern blot analysis was performed using strand-specific sRNA probes radiolabeled with [α-^{32}P] UTP. All blots included RNA samples from control (uninfected HMECs) and those infected for 3h and 24h with *R. prowazekii*. The blot for 6S RNA only included RNA isolated from HMECS that were either left uninfected or processed at 24 h post-infection.

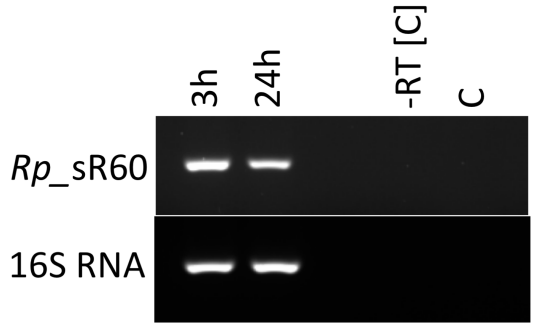


Figure 14. RT-PCR analysis.
A representative image of the RT-PCR analysis for *Rp*_sR60 is shown. All blots included cDNA samples from control (uninfected HMECs) and those infected for 3h and 24h with *R. prowazekii*.

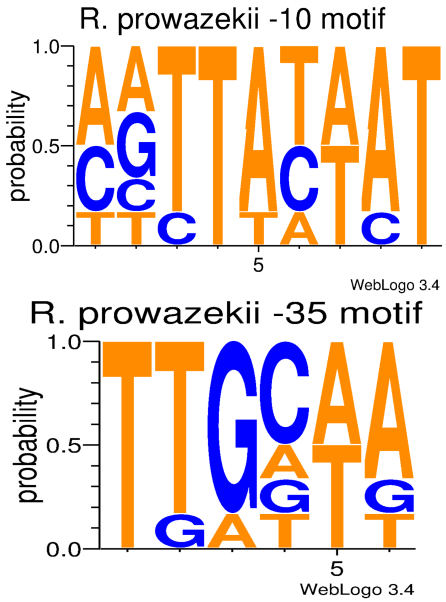


Figure 15. sRNA promoter frequencies.

Conservation diagrams illustrating the probability of a nucleotide in a specific promoter motif position based on the confirmed sRNAs. Top panel shows the -10 promoter motif, while bottom panel represents the -35 promoter motif. The predicted -10 motifs are similar to the *E. coli* consensus sequence (TATAAT). On the other hand, the -35 motifs differ in comparison to the *E. coli* consensus sequence (TTGACA).

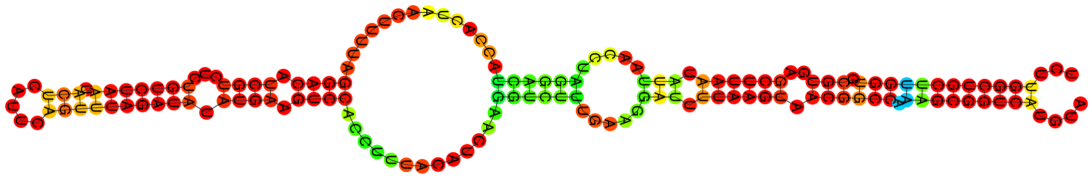


Figure 16. Predicted secondary structure of *R. prowazekii* 6S RNA.
Predictive analysis of the secondary structures of identified sRNAs was accomplished using RNA-fold. The color palette represents base-pairing probability from 0 to 1 (purple to red). This secondary structure of *R. prowazekii* 6S RNA resembles the previously reported consensus secondary structure of bacterial 6S RNA.

Table 12. Identified transcription start sites determined by RLM-RACE and associated promoter motifs.

sRNA	-10 Box	-10 Position	-35 Box	-35 Position	Start Site
17	TGCTTTTAT	-93 bp	TTGCAA	-114 bp	105,761
34	CTTTATAAT	-45 bp	TTGCTA	-69 bp	326,320
47	AATTAAAAT	-65 bp	TGGATA	-89 bp	481,512
60	AGTTACTAT	-43 bp	TTGGTG	-65 bp	844,442
67	AATTATAAT	-36 bp	TTGCAT	-56 bp	998,424
6S	CCTTACTCT	-77 bp	TTATAA	-94 bp	934,943

Using the RLM-RACE sequencing data, the -10 motif and the -35 motif for the σ^{70} promoter were predicted using BPROM. The table lists the predicted -10 nucleotide sequence, the -10 start position upstream from the transcription start site, the -35 nucleotide sequence, and the -35 start position upstream from the transcription start site for the confirmed rickettsial sRNAs. The "Start Site" is the transcription start site determined by RLM-RACE.

Table 13. Prediction of target genes using TargetRNA2 and IntaRNA.

sRNA candidate name	Total number of target genes predicted by TargetRNA2	Total number of target genes predicted by CopraRNA	Number of common targets predicted by both programs	List of target genes predicted by both programs	Protein name
*Rp*_sR17	19	44	1	H375_290	Protein translocase subunit SecD
*Rp*_sR60	23	40	6	H375_7590	3-oxoacyl-[acyl-carrier-protein] synthase 2
				H375_4980	SSU ribosomal protein S7p (S5e
				H375_3020	Putative TolA protein
				H375_1850	Phosphatidate cytidylyltransferase
				H375_8280	Putative cytochrome c-type biogenesis protein related to CcmF
				H375_8880	LSU ribosomal protein L6p (L9e)
*Rp*_sR67	30	37	7	H375_8490	Methionine--tRNA ligase
				H375_6610	Alanine--tRNA ligase
				H375_4810	Hypothetical protein H375
				H375_2550	Reductase
				H375_1110	tRNA pseudouridine synthase B H375
				H375_3730	Uncharacterized protein RP244 of Rickettsia
				H375_6090	ATP synthase subunit c H375

Listed are the common set of targets predicted by both TargetRNA2 and IntaRNA. The total number of targets predicted by TargetRNA2 is shown next to the total number of targets predicted by IntaRNA. The number of common targets is documented next to a list of those common targets and their functions.

Chapter 5 Define the expression profile of novel sRNAs between the virulent strains of *R. prowazekii* and their mechanistic role during host-pathogen and vector-pathogen interactions

INTRODUCTION.

Despite being essentially ignored for nearly 40 years, sRNAs are now viewed as a critical-regulators for bacterial gene expression (Pichon and Felden, 2008). Intense focus into these RNA transcripts has yielded their reports in practically all Eubacteria families as well as some Archaea families (Updegrove *et al.*, 2015). In the previous chapters, I showed that *Rickettsia* encode and express small RNAs during infection of human microvascular endothelial cells (Schroeder *et al.*, 2015; Schroeder *et al.*, 2016). These studies excluded the sRNA expression during infection of an arthropod vector, which represents a large portion of the rickettsial life cycle.

Traditionally, *Pediculus humanus corporis* (human body louse) has been considered the vector for *Rickettsia prowazekii*. However, past reports have described isolation of the pathogen from non-traditional sources, such as *Amblyomma variegatum, Hyalomma marginatum, and Hyalomma truncatum* ticks from Ethiopia and fleas (Reiss-Gutfreund, 1966). Considered to be controversial at the time as repeated searches in ticks failed to find *R. prowazekii* (Philip *et al.*, 1966), *Amblyomma cajennense* ticks from Mexico were found to harbor the bacterium (Medina-Sanchez *et al.*, 2005). Few reports have investigated bacterial sRNAs expression during infection of the arthropod vector. One such study examined the sRNA expression profile of *Wolbachia pipientis,* an α-

proteobacteria distantly related to *Rickettsia*, during infection of *Drosophila melanogaster* (commonly known as the fruit fly). The authors noted that *Wolbachia*'s putative sRNAs showed a differential expression profile based on the fly's sex and the specific tissues from which the pathogen was isolated (Woolfit *et al.*, 2015). Nonetheless, this species is not known to naturally infect humans.

While it is vital to understand sRNA expression during human infection in order to yield a greater understanding of rickettsial virulence and pathogenesis, it is equally importantly to characterize sRNA expression during infection of an arthropod vector. In this way, novel arthropod-pathogen interactions may be identified, shedding light on bacterial maintenance and its life cycle. To address this knowledge gap, I infected *Amblyomma americanum* tick cell line (AAE2) with *Rickettsia prowazekii* strain Breinl and identified additional 96 novel *trans*-acting and *cis*-acting sRNAs. In addition, I report the confirmed interaction between *Rp*_sR60, a previously confirmed sRNA, and H375_0420, and further characterizing its expression profile in the human host and the arthropod vector.

RESULTS

Rp_sR60 sRNA Target Identification

Cis-acting sRNAs primarily regulate expression of the gene located on the sense strand. On the other hand, *trans*-acting sRNAs regulate the expression of gene targets located elsewhere in the genome through partial base pairing that either represses or stimulates translation. For this reason, the determination of functional roles for *trans*-acting sRNAs initially required a bioinformatics approach to predict and analyze potential targets. To accomplish this, I employed the web-based program CopraRNA to

predict a set of targets for *Rp*_sR60. Initially in Aim #2, I predicted targets for this sRNA using a combination of TargetRNA2 and IntaRNA. However, it has been reported that with the addition of phylogenetic conservation into an RNA target prediction program, as was done with CopraRNA, a large gain of target accuracy is had (Pain *et al.*, 2015). This gain resulted in a lower rate of false positives for CopraRNA predictions when compared to TargetRNA2 and IntaRNA. As such, a new set of predictions was created for *Rp*_sR60 using CopraRNA. Using the basic parameters, the algorithm predicted a total of 53 targets with a $p \leq 0.05$. A detailed analysis revealed that 85% of the predicted targets had assigned function. The *R. prowazekii* genome annotation includes a high percentage of either hypothetical or uncharacterized proteins; thus, it is reasonable to speculate that *Rp*_sR60 has some regulatory influence in *R. prowazekii*. *Rp*_sR60 regulated pathways macromolecule biosynthesis, transport, ribosomal RNA formation, and other cellular and metabolic functions. Furthermore, *Rp*_sR60 has also been shown to regulate VirD4, an ATP hydrolysis protein that is essential to energy production for the Type IV Secretion System. The predicted target genes regulated by *Rp*_sR60 are listed in Additional File 6.

Experimental Validation of Rp_sR60-mRNA Target Interaction.

In order to confirm targets predicted by CopraRNA, electrophoretic mobility shift assays (EMSA) were carried out using *in vitro* transcription-generated [α-^{32}P] UTP labeled sRNA transcripts and unlabeled target transcripts mixed in a 2:1 molar ratio. Mixtures were incubated at 30°C or 37°C to simulate arthropod and human host temperatures, respectively. Only ten target predictions (marked with star in Additional

File 7) were analyzed for interaction as these targets exhibited the strongest seed region energies. Namely, these were H375_8740, H375_7670, H375_1650, H375_2860, H375_7680, H375_1040, H375_6080, H375_8760, H375_0420, and H375_2090, in order of the strongest p-value. Interestingly, *Rp*_sR60 formed a stable RNA duplex at both 30°C and 37°C (n=3 for both temperatures) with H375_0420, excinuclease ABC subunit C (*uvrC*). The RNA complex was absent during cold competition, thus indicating a true positive interaction. No interactions were observed with the other selected targets (Figure 17).

In Vitro Expression Profile for Rickettsial H375_0420 and Rp_sR60 During Infection of HMECs and AAE2 Cell Lines

Although the human body louse is considered the primary natural vector, *R. prowazekii* has been isolated from *Amblyomma* ticks in Mexico and Ethiopia, indicating that body lice are not the only natural arthropod vector (Philip *et al.*, 1966; Medina-Sanchez *et al.*, 2005). Therefore, I used the well-established AAE2 tick cell line created from *Amblyomma americanum* nymphs as a model of arthropod infection. The expression of *Rp*_sR60 and H375_0420 was assessed at 0.5 h, which served as the baseline, 3 h, and 24 h post-infection by RT-PCR using target specific primers and 16S rRNA as the endogenous control. Interestingly, expression of *Rp*_sR60 and H375_0420 was found in HMECs and AAE2 cells, indicating a potential function during both arthropod and human infections. Further, in AAE2 infection, both H375_0420 and *Rp*_sR60 kept a steady expression when compared to their respective 0.5 h baseline (Figure 18). There were no significant differences in expression between *Rp*_sR60 and

H375_0420 at 3 h and 24 h, respectively. However, for HMEC infection, both *Rp*_sR60 and H375_0420 showed a significant decrease ($p<0.01$) when compared to their respective baselines (Figure 19). When analyzed between *Rp*_sR60 and H375_0420 at 3 h and 24 h, respectively, there were no significant differences in expression. These results suggest that *Rp*_sR60 may positively regulate H375_0420.

RNA Sequencing of Rickettsia prowazekii Infected AAE2 Cells

Using the web-based SIPHT interface, I predicted 26 candidate *trans*-acting small RNAs encoded in the *R. prowazekii* strain Breinl genome. Using RNA-seq to verify their existence during HMEC infection, 12 candidates were found to have a Mean Expression Value (MEV) of >1.5. RT-PCR confirmed the expression of six potential sRNAs (Additional File 1 and Additional File 3) (Schroeder *et al.*, 2015). As the SIPHT program only predicts candidate trans-acting sRNAs based on specific parameters and factors, a deeper analysis of the RNA-seq data demonstrated an additional 35 *trans*-acting and 23 *cis*-acting sRNAs at either 3 h or 24 h post-infection (Additional File 3) (Schroeder *et al.*, 2016). To investigate expression of novel sRNAs during arthropod infection, I performed next generation sequencing to identify and catalogue arthropod-specific sRNAs encoded in *R. prowazekii* strain Breinl.

Serving as a model for late infection, RNA sequencing of AAE2 and HMECs infected with *R. prowazekii* strain Breinl for 24 h resulted in approximately 91.15 million and 46.1 million total reads, respectively. Despite treatment methods to enrich the bacterial transcripts, the majority of total reads consisted of eukaryotic transcripts such as mitochondrial transcripts, structural transcripts, and long and small non-coding RNAs.

Overall, the AAE2 sample provided 2,982,823 reads to the bacterial genome, while the HMEC sample consisted of 1,255,181 bacterial reads. Prior to sequencing, half of each RNA sample was subjected to TEX treatment to remove any processed RNA transcripts, of which 333,431 mapped to the HMEC-infected *R. prowazekii* genome and 256,084 reads mapped to AAE2-infected *R. prowazekii* genome. While initially appearing limited, the percentage of reads mapping to the bacterial genome correlated with previous next generation sequencing conducted by the laboratory (Schroeder *et al.*, 2016). In addition, the eukaryotic-to-bacteria ratio obtained in the RNA-seq data strongly matched a finding that obligate intracellular bacterial genomes only constitute 2-5% of extracted total RNA despite efficient enrichment protocols. Further, only 5% of the extracted bacterial RNA is composed of mRNAs and sRNAs, with the remaining 95% mapping to rRNAs and tRNAs (Westermann *et al.*, 2012).

An RNA-seq based search of the AAE2-infected transcriptome resulted in the identification of an additional 93 *trans*-acting (intergenic) and/or *cis*-acting (antisense) sRNAs at 24 h. *Cis*-acting composed 72% of the novel sRNAs. Generally speaking, sRNAs range from 50 to 500 bp in length, although there are exceptions, such as RNAIII encoded by *Staphylococcus aureus* at 514 bp (Boisset *et al.*, 2007), a 581 bp anti-sense sRNA in *Pseudomonas aeruginosa* (Gómez-Lozano *et al.*, 2014), and even a 1.2kb anti-sense RNA encoded by *Salmonella enterica,* termed AmgR (Lee and Groisman, 2010). In fact, a recent examination of the sRNA profile in *Burkholderia pseudomallei* identified nine sRNAs over 500 bp, with the longest sRNA at 750 bp (Ooi *et al.*, 2013). *Rp*_sR152, a *cis*-acting sRNA antisense to H375_8370, was identified as the smallest sRNA at 148 bp in length. Interestingly, 11 sRNAs (*Rp*_sR79, *Rp*_sR92, *Rp*_sR93, *Rp*_sR94,

*Rp*_sR114, *Rp*_sR120, *Rp*_sR128, *Rp*_sR130, *Rp*_sR135, *Rp*_sR143, and *Rp*_sR154) were identified that ranged from 516 bp to 844 bp in length, 8 of which were classified as *cis*-acting. No strand bias was observed with 47% of sRNAs on the leading strand and 53% on the lagging strand. Further, no significant differences were observed based on the type of sRNA as 38% *trans*-acting and 50% *cis*-acting sRNAs were located on the leading strand. When examining the transcription direction of the *trans*-acting sRNAs, I compared the *trans*-acting sRNA and the upstream gene orientations. I found 6 of the 26 *trans*-acting sRNAs were encoded in the same orientation as their upstream genes (*Rp*_sR71, *Rp*_sR74, *Rp*_sR75, *Rp*_sR129, *Rp*_sR139, and *Rp*_sR155). In regards to the context of the cis-acting sRNAs, 66 sRNAs were found to be directly anti-sense to only one ORF, as generally expected. Interestingly, however, *Rp*_sR124 overlaps the 3' end of H375_7030, a 3-polyprenyl-4-hydroxybenzoate carboxylase, and the 5' end of H375_7040, polyhydroxyalkanoic acid synthase, a space of approximately 50 bp.

qRT-PCR of Novel Differentially Expressed sRNAs

As previously reported, bacterial sRNA expression can differ between tissues and organ systems within an infected host as well as between genders of a species (Woolfit *et al.*, 2015). It stands to reason that sRNA expression can vary depending on the host (human vs arthropod). Therefore, qRT-PCR was employed to assess the differences in expression for five novel sRNAs identified in the AAE2 model to that of the human model. Rickettsial sRNAs were assessed at 0.5 h, which served as the baseline, 3 h, and 24 h post-infection using sRNA specific primers and 16S rRNA as the endogenous control (Table 3). *Rp*_sR76 and *Rp*_sR86 demonstrated significantly higher levels of

expression at 3 h when compared to their respective HMEC expression ($p \leq 0.001$), but failed to show any significance differences at 24 h (Figure 20 and Figure 22, respectively). Likewise, *Rp*_sR83 and *Rp*_sR159 also showed a significantly higher expression at 3 h and no significant differences at 24 h when compared to the respective HMEC expression levels ($p \leq 0.05$) (Figure 21 and Figure 23). Interestingly, there was no evidence of any significant differences between HMEC and AAE2 for *Rp*_sR63 (data not shown).

DISCUSSION

It is now widely accepted that sRNAs are critical post-transcriptional regulators of bacterial gene expression (Pichon and Felden, 2008). The availability of bacterial genomes and more cost-effective methods for identification has allowed intensive studying and characterization of these noncoding transcripts. While initial screening methods for sRNAs included computational searches, today both bioinformatics and next generation sequencing are the primary methods of detecting small RNAs. However, identification still remains difficult due to the poor understanding of recurring nucleotide motifs (Pichon and Felden, 2008). Further, sRNAs are commonly expressed in specific environmental situations, such as iron limitation, cell stress, or nutrient starvation, making them difficult to detect globally (Updegrove *et al.*, 2015). Further, searches for sRNA repertoires in intracellular bacteria, such as *Rickettsia*, remain particularly challenging due to low RNA yields (Khandige *et al.*, 2015). The total number of encoded sRNAs within any bacterial species remains unknown, although it is assumed to be

around a few hundred as opposed to the few thousands as once thought (Gottesman and Storz, 2011).

Recently, *R. prowazekii* was shown, using next generation sequencing, to express a repertoire of both *cis*-acting and *trans*-acting sRNAs during infection of human microvascular endothelial cells (HMECs) (Schroeder *et al.*, 2015; Schroeder *et al.*, 2016). *Rp*_sR60 was selected due to its orientation in the genome and the number of mapped reads for an mRNA target computational search. Of the ten targets predicted using CopraRNA, only H375_0420, a known ORF to encode for UvrC, displayed a stable RNA complex during electrophoretic mobility shift assays, verifying its ability to interact with *Rp*_sR60. As *trans*-acting sRNAs frequently require a chaperone protein to facilitate binding of the target and the sRNA, of which there are currently none described with this function in *Rickettsia,* it must be noted that *Rp*_sR60 may still in fact bind with the other nine predicted targets (Östberg *et al.*, 2004). Analysis of the predicted interaction showed that *Rp*_sR60 interacts with H375_0420 approximately 150 bp upstream of the start codon (genomic location 49914 and 49936) (Figure 24). UvrC is a dual excinuclease with the function to nick a damaged DNA strand at both 5' and 3' of the lesion site (Van Houten and Kad, 2014). After DNA damage, particularly UV damage that causes dimerization of specific nucleotides, the prokaryotic nucleotide excision repair system is initiated using one of two mechanisms. The first, known as the global genome repair, is initiated when a scanning $UvrA_2UvrB_2$ complex detects a distortion in the DNA strand and recruits UvrC. The other method, transcription-coupled repair (TCR), occurs during transcription when an RNA polymerase is blocked due to a lesion and is pushed off by Mfd, a TCR factor, which recruits $UvrA_2$ to the site. At this point,

both systems proceed the same as UvrB guides UvrC to the lesion site for an incision that is then repaired by UvrD and RNA Polymerase I (Van Houten and Kad, 2014). To prevent tuberculosis disease progression caused by *Mycobacterium tuberculosis,* targeting UvrC has been suggested since it was reported to enhance pathogenicity (Parulekar *et al.*, 2013). Further, the SOS response system induces *uvrA* and *uvrB,* which indirectly induces *uvrC* (Janion, 2008). Inactivation of either UvrA, UvrB, or UvrC would sabotage the nucleotide excision repair and TCR systems, opening the pathogen to deleterious DNA damage. A BLAST analysis reveals UvrA, UvrB, and UvrC to be highly conserved within the genus *Rickettsia* (data not shown).

To define the transcriptional catalogue for *R. prowazekii* during the arthropod stage of its life cycle, I used next generation sequencing as a global high throughput method to identify potential small RNAs. RNA sequencing improves the extent of the transcriptome and the accuracy for genome annotations (Sharma and Vogel, 2014). The search, based on the transcript orientation and genome location, the orientation of the up- and down-stream genes, and the number of mapped reads, identified an additional 93 small RNAs expressed at 24 h post-infection that can be included with the already identified 70 sRNAs at 3 h and 24 h post-infection in HMECs. Five of these were selected for further characterization during host- and vector-pathogen interactions. Namely, *Rp*_sR76, *Rp*_sR83, *Rp*_sR86, and *Rp*_sR159 were found through qRT-PCR to have significantly higher levels of expression at 3 h post-infection in arthropod cells, suggesting that these sRNAs are niche-dependent. However, *Rp*_sR63 showed no significant differences in expression at 3 h or 24 h post-infection. These results were expected as the majority of sRNAs, especially *trans*-acting, are expressed under specific

environmental conditions (Woolfit *et al.*, 2015). A recently identified sRNA in *Wolbachia,* termed *ncrwmel02*, was reportedly expressed at two and seven times higher levels in *W. pipientis* strain wMel and strain wAu, respectively, when compared to strains wMelCs and wMelPop. Further, *ncrwmel02* in strain wMel was expressed at significantly higher levels in male *Drosophila* abdomens than female abdomens and was upregulated in testes compared to ovaries, suggesting tissue-specific and host sex-specific regulation (Woolfit *et al.*, 2015). In the *β-proteobacteria Burkholderia thailandensis*, microarray analysis identified differential expression of 38 novel and 2 conserved sRNAs in response to 54 environmental conditions (Stubben *et al.*, 2014). Likewise, a comprehensive transcriptomic analysis revealed differential expression of sRNAs encoded in *Burkholderia pseudomallei* under 82 environmental conditions (Ooi *et al.*, 2013).

The majority of sRNAs described in Gram-negative bacteria have been classified as *trans*-acting and require the assistance of Hfq, a chaperone protein that facilitates the binding of the *trans*-acting sRNA and the mRNA target (Khandige *et al.*, 2015). However, approximately 50% of bacteria do not encode for Hfq, including *Rickettsia* (Östberg *et al.*, 2004; Dugar *et al.*, 2013). Despite *Rickettsia* being a Gram-negative organism, approximately 25% of the identified sRNAs during AAE2 infection were categorized as *trans*-acting. This is contradictory to the current school of thought that Gram-negative organisms rely mostly on *trans*-acting sRNAs, while Gram-positive organisms rely primarily on *cis*-acting elements (Waters and Storz, 2009; Lasa *et al.*, 2012). It should be noted that research direction and the previous lack of technologies to detect *cis*-acting sRNAs may skew these assumptions (Waters and Storz, 2009). In a

transcriptomic search of *B. pseudomallei,* approximately 10% of genes were found to have antisense transcription (Ooi *et al.*, 2013). Likewise, a global analysis of *Helicobacter pylori* found at least one antisense transcription start site for nearly 46% of all ORFs and 28% of tRNAs (Sharma *et al.*, 2010). For *Staphylococcus aureus*, a Gram-positive bacterium, antisense transcription was reported for nearly 50% of its ORFs (Lasa *et al.*, 2011). Combining the AAE2 novel sRNAs and the previously identified sRNAs during HMEC infection, approximately 44% have been classified as *trans*-acting and 55% classified as *cis*-acting, bringing this ratio closer to other reported organisms.

Originating from the anti-sense of a coding gene, *cis*-acting sRNAs bind to their complementary coding transcript presumably resulting in post-transcriptional regulation. In this study, I identified 67 *cis*-acting sRNAs antisense to key ORFs encoding for structural proteins, transporters, membrane lipoproteins, and metabolic pathways. Namely, *Rp*_sR101 and *Rp*_sR102 are antisense to H375_5210, a VirB6 paralog, and H375_5270, the ATPase VirB4, respectively. The Outer Membrane Protein B, H375_8270, harbors the cis-acting *Rp*_sR148. Further, H375_1620, an ATP-dependent helicase encoding UvrD, is hypothetically regulated via *Rp*_sR77. Spanning multiple membranes, the Type IV Secretion System (T4SS) is a complex multi-protein transporter in many Gram-negative organisms. Acting similar to a hypodermic syringe, the system allows bacteria to inject a variety of virulence factors into eukaryotic hosts. For this reason, it has garnered significant attention in the field of rickettsiology (Gillespie *et al.*, 2009; Gillespie *et al.*, 2015). The T4SS is composed of at least 12 proteins, many of which have multiple paralogs with unknown individual functions. The VirB6 component comprises the T4SS inner channel and is essential to substrate transfer, making it the

most divergent of the VirB proteins with at least five paralogs (Gillespie *et al.*, 2009). Likewise, the VirB4 protein is an integral part of the T4SS due to its ATPase activity, which provides the required energy for operation. Outer membrane protein B (OmpB), also known as *sca5*, is an abundant protein expressed on the surface of all species of *Rickettsia*. Its function, among a few other rickettsial outer membrane proteins, enables *Rickettsia* to bind to and invade the eukaryotic host cell. Antibodies directed against OmpB have been reported to protect mice from lethal doses of *Rickettsia* (Chan *et al.*, 2009). As alluded to above in the nucleotide excision repair and the transcription coupled repair description, after UvrC incises 5' and 3' of a DNA lesion, UvrD excises the dimerized nucleotides for final repair by RNA Polymerase I (Van Houten and Kad, 2014).

Further investigation into rickettsial sRNAs will likely lead the identification of novel sRNAs in all species of *Rickettsia*. In addition, different environmental conditions that *Rickettsia* may experience, such as overwintering in ticks, will lend itself to the identification of additional sRNAs. Determination as to whether these sRNAs up-regulate or down-regulate gene expression will shed new insights into rickettsial virulence. As the field of rickettsial pathogenesis evolves, a deeper understanding for rickettsial virulence factors and rickettsial transcriptomics will emerge. This study provided the very first glimpse of rickettsial small RNAs expressed in the arthropod vector and validated that H375_0420 interacts with *Rp*_sR60.

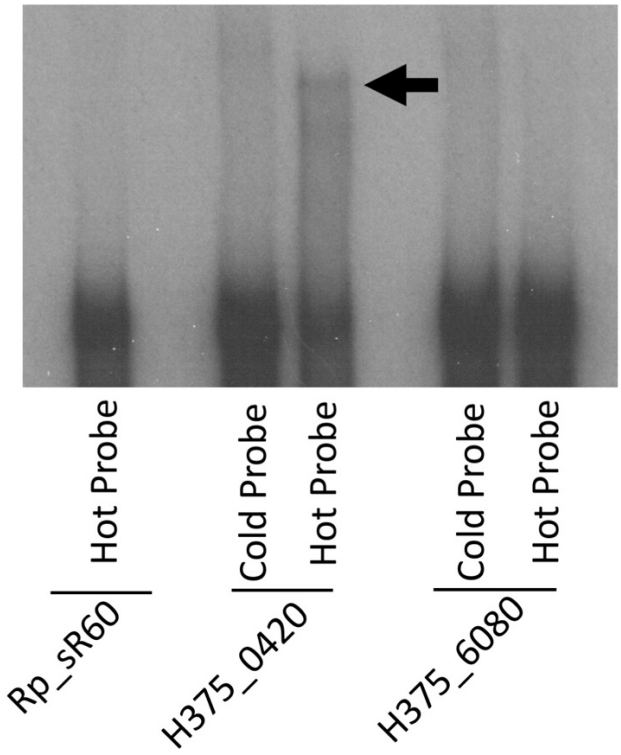


Figure 17. *Rp*_sR60 EMSA with CopraRNA predicted targets.
*Rp*_sR60 electrophoretic mobility shift assays (EMSA) were performed on selected CopraRNA predicted mRNA targets. [α-^{32}P] UTP labeled sRNA transcripts and unlabeled target transcripts were created through *in vitro* transcription and mixed at a 2:1 molar ratio followed by incubation at 30°C or 37°C. Cold competitions were carried out using unlabeled sRNA transcripts mixed with unlabeled RNA targets chased by radiolabeled sRNA transcripts. A stable RNA complex was observed for H375_0420 for the "hot probe" lane only. No RNA complexes were observed in the cold competition lane, nor were any observed with the other predicted targets (e.g. H375_6080).

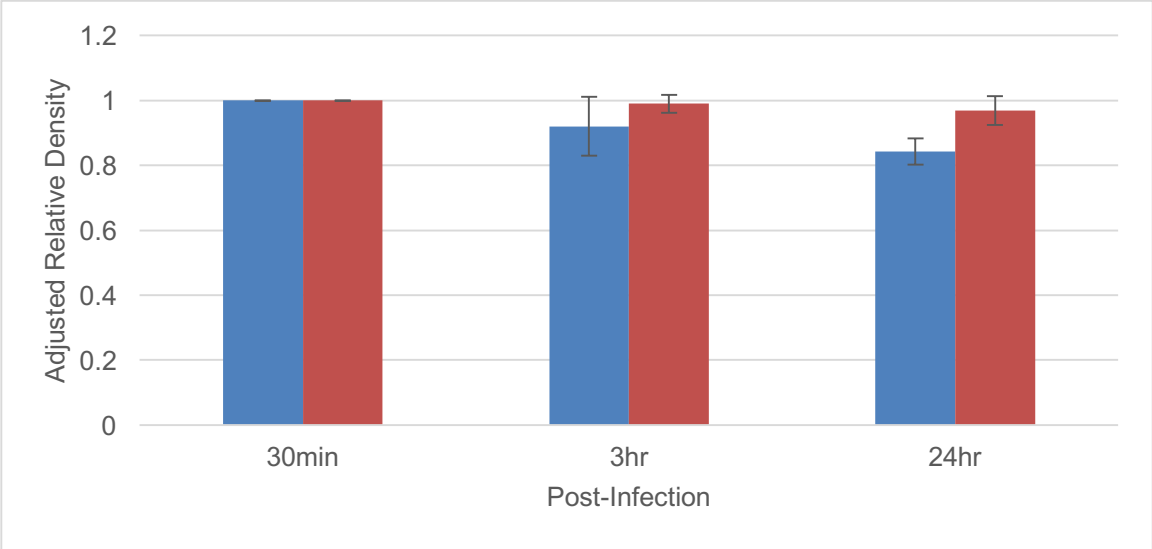


Figure 18. *Rp*_sR60 and H375_0420 expression using AAE2 model of infection.
Confluent monolayer of AAE2 cells were infected with *Rickettsia prowazekii* for 0.5 h, 3 h, and 24 h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for reverse transcriptase PCR (n=3). *Rp*_sR60 (blue) and H375_0420 (red) expression were normalized to 16S rRNA with 30 minutes as baseline gene expression. Relative densities were measured using ImageJ. No significant difference was observed between baseline expression or between targets at a specific time point.

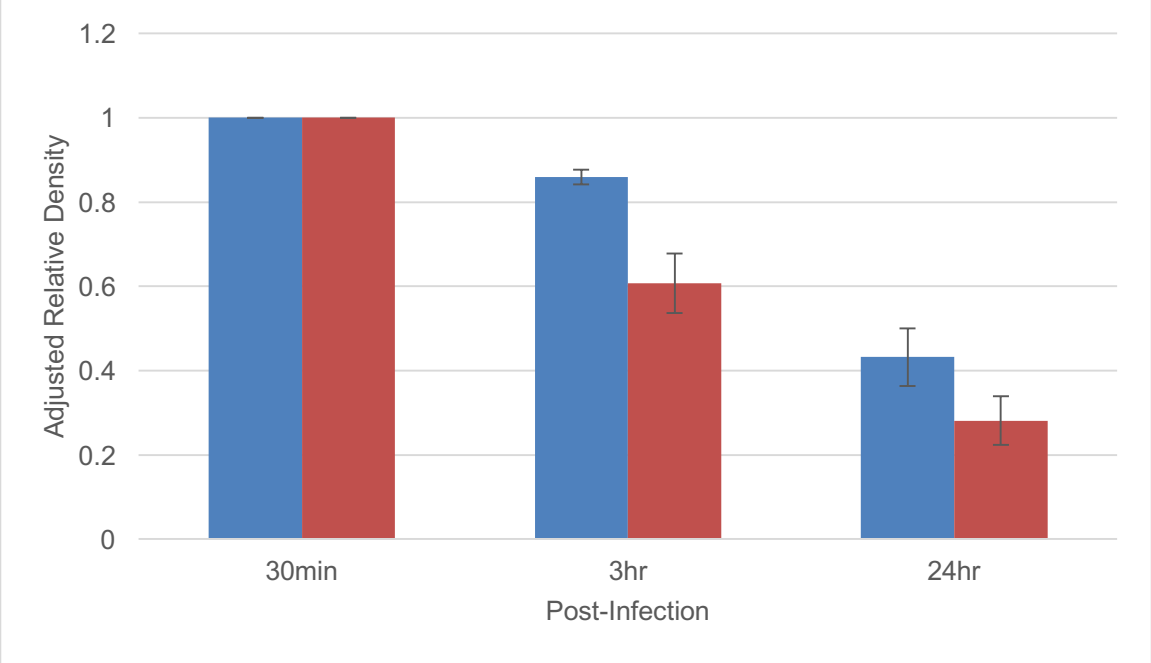


Figure 19. *Rp*_sR60 and H375_0420 expression using HMEC model of infection.
Confluent monolayer of HMECs were infected with *Rickettsia prowazekii* for 0.5 h, 3 h, and 24 h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for reverse transcriptase PCR (n=3). *Rp*_sR60 (blue) and H375_0420 (red) expression were normalized to 16S rRNA with 30 minutes as baseline gene expression. Relative densities were measured using ImageJ. A significant decrease was observed between 3 h and 24 h when compared to baseline. However, no significant difference was observed between targets at a specific time points.

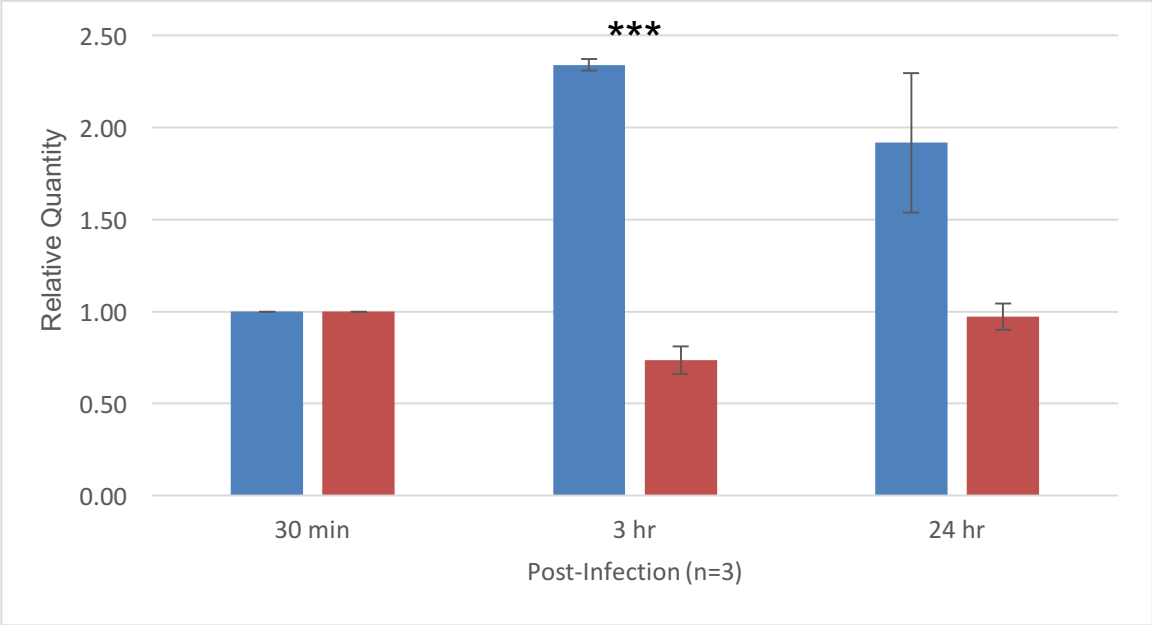


Figure 20. *Rp*_sR76 expression using HMEC and AAE2 models of infection.
Confluent monolayer of HMECs or AAE2 were infected with *Rickettsia prowazekii* for 0.5 h, 3 h, and 24 h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for RT-PCR (n=3). AAE2 (blue) and HMEC (red) expression were normalized to 16S rRNA. A significantly higher level ($p \leq 0.001$) of expression was observed for AAE2 at 3 h when compared to HMEC. No significance was observed for 24 h.

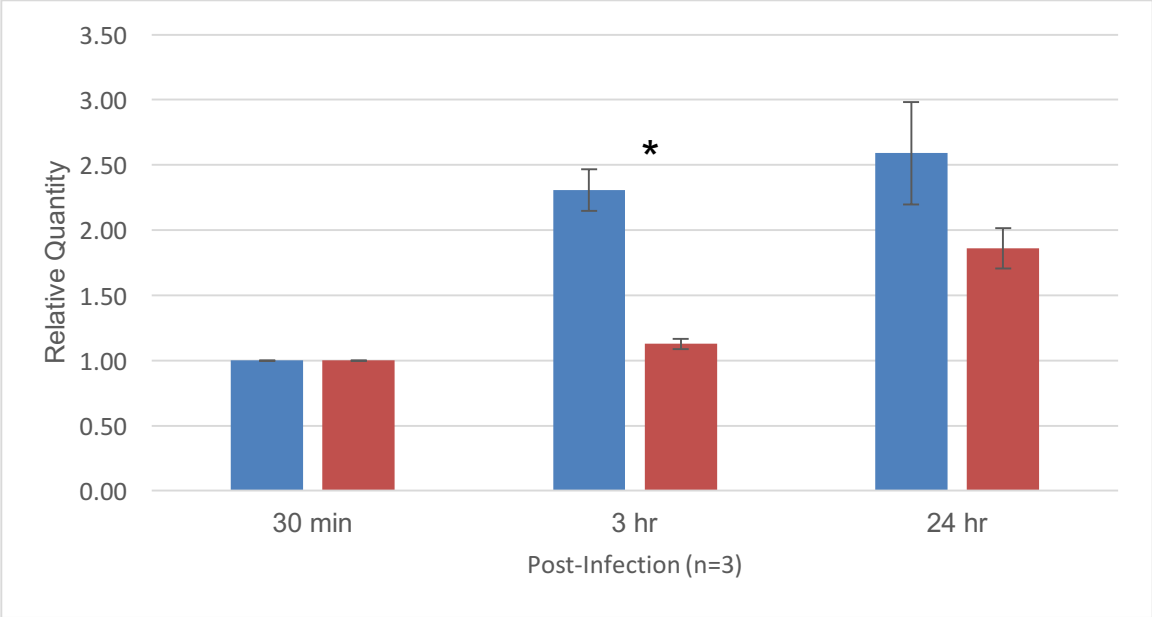


Figure 21. *Rp*_sR83 expression using HMEC and AAE2 models of infection.
Confluent monolayer of HMECs or AAE2 were infected with *Rickettsia prowazekii* for 0.5 h, 3 h, and 24 h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for RT-PCR (n=3). AAE2 (blue) and HMEC (red) expression were normalized to 16S rRNA. A significantly higher level ($p \leq 0.05$) of expression was observed for AAE2 at 3 h when compared to HMEC. No significant differences were observed at 24 h.

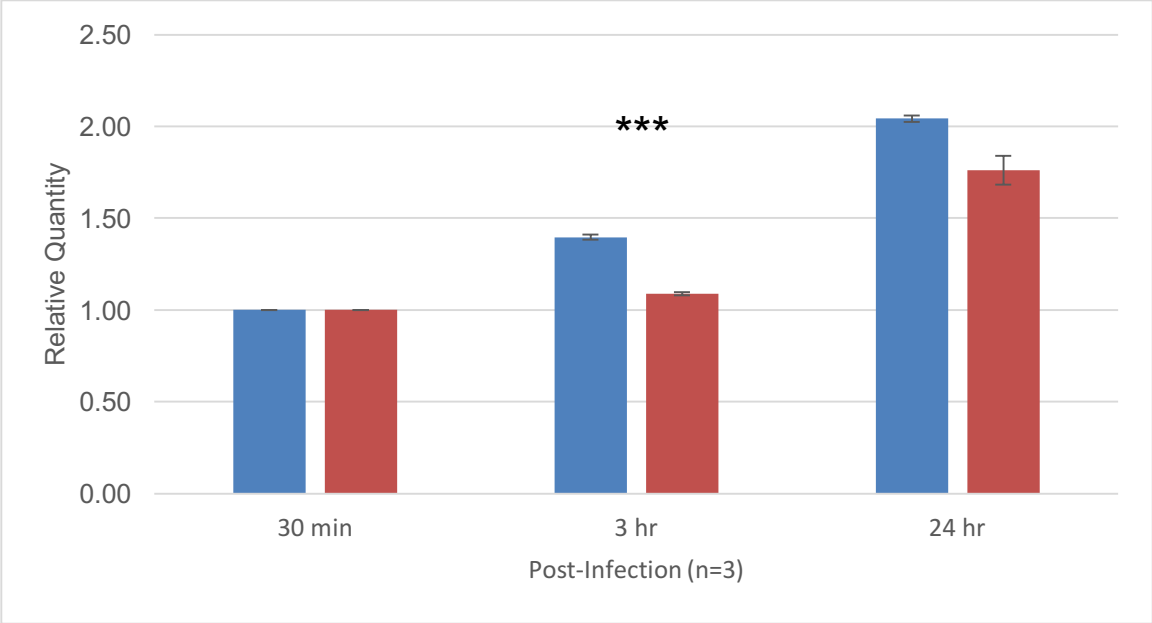


Figure 22. *Rp*_sR86 expression using HMEC and AAE2 models of infection.
Confluent monolayer of HMECs or AAE2 were infected with *Rickettsia prowazekii* for 0.5 h, 3 h, and 24 h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for RT-PCR (n=3). AAE2 (blue) and HMEC (red) expression were normalized to 16S rRNA. A significantly higher level ($p \leq 0.05$) of expression was observed for AAE2 at 3 h when compared to HMEC. No significant differences were observed at 24 h.

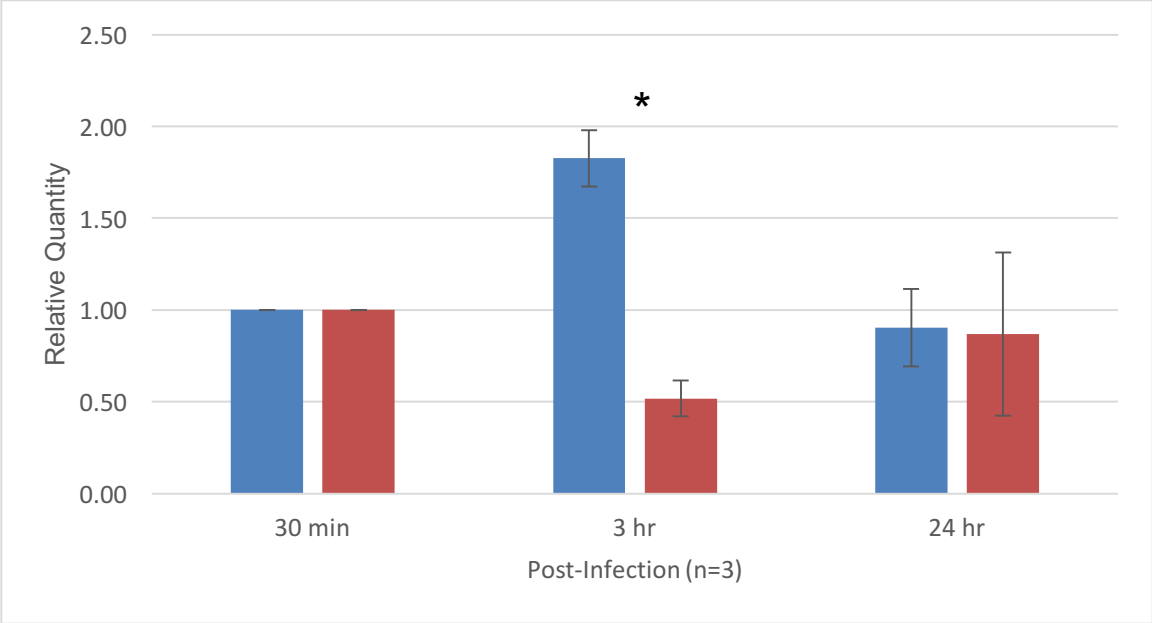


Figure 23. *Rp*_sR159 expression using HMEC and AAE2 models of infection.
Confluent monolayer of HMECs or AAE2 were infected with *Rickettsia prowazekii* for 0.5 h, 3 h, and 24 h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for RT-PCR (n=3). AAE2 (blue) and HMEC (red) expression were normalized to 16S rRNA. A significantly higher level ($p \leq 0.05$) of expression was observed for AAE2 at 3 h when compared to HMEC. No significant differences were observed for 24 h.

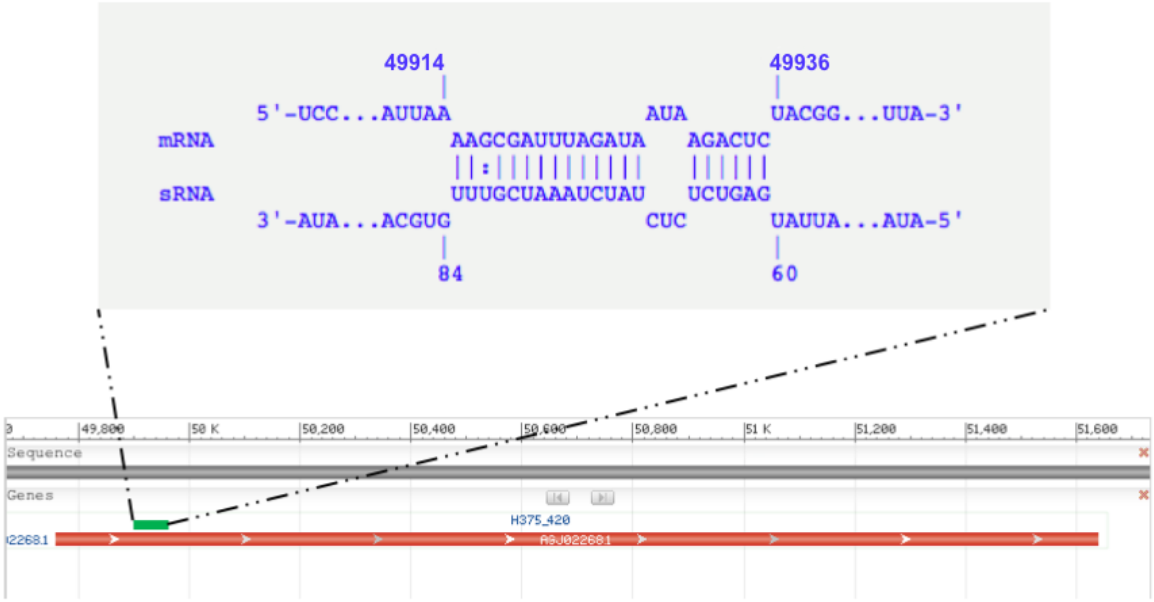


Figure 24. *Rp*_sR60 interaction with H375_0420.

*Rp*_sR60, which was validated previously to be expressed in human microvascular endothelial cells, binds with the mRNA target H375_0420 as predicted by CopraRNA. A stable RNA complex was observed in electrophoretic mobility shift assays (EMSAs).

Chapter 6 Summary

Rickettsia are medically relevant obligately intracellular bacterial pathogens that cause severe disease. As the causative agents of epidemic typhus and Rocky Mountain spotted fever, these neglected pathogens have had significant impacts on society with mortality rates reaching upwards of 50%. Clinical diagnosis of rickettsial infections is often difficult owing to common flu-like symptoms such as nausea, headache, fever and vomiting, which are invoked by several other bacterial and viral pathogens. For this reason, antibiotic therapies are often delayed or inappropriately administered. The central hypothesis of this work is that since genomic reduction, the main driving force shaping rickettsial genomes, results in maintaining only a minimal set of genes required for pathogenesis and the obligate intracellular lifestyle, the *Rickettsia* genome encodes sRNAs required for regulation of its transcriptome, and differentially expresses these sRNAs leading to altered virulence and adaptation depending on the host niche.

The first aim analyzed the potential for sRNAs in a number of *Rickettsia* species using a combination of bioinformatics and *in vitro* methods. Using the SIPHT graphics user input to analyze the intergenic regions of 16 rickettsial strains, which represented 13 species, I predicted a potential of 1,700 small RNAs to be present. Selecting *R. prowazekii* and *R. typhi* from the typhus group and *R. rickettsii* and *R. conorii* from the spotted fever group, each of the predicted sRNAs were further characterized using BPROM and TransTermHP for promoters and Rho-independent terminators, respectively. Each of the sRNAs analyzed had strong σ70 promoters within 150 bp upstream, thus indicating the potential for transcriptional activity. Due to its medical

importance and potential for bioterrorism, *R. prowazekii* was chosen for experimental validation of the 26 sRNA candidates using next generation sequencing to analyze the bacterial transcriptome during infection of human microvascular endothelial cells (HMECs) at 3 h and 24 h post-infection. Six sRNA candidates were further verified using reverse transcriptase PCR. These results yielded clear evidence for sRNA expression during *R. prowazekii* infection of HMECs.

The second aim of this project identified novel *R. prowazekii* sRNAs and validated well-conserved bacterial sRNAs. Next generation sequencing of human microvascular endothelial cells infected with *R. prowazekii* exposed the expression of 35 *trans*- and 23 *cis*-acting sRNAs. Further validation through Northern blot or RT-PCR analyses confirmed the expression of two *trans*-acting (*Rp*_sR17 and *Rp*_sR60) and one *cis*-acting (*Rp*_sR47) novel sRNAs. In addition, Northern blot analysis verified the expression of four well-conserved sRNAs (RNaseP_bact_a, α-tmRNA, 4.5S RNA, 6S RNA). Selecting five novel rickettsial sRNAs and 6S RNA, the transcriptional start sites were determined using 5' RLM-RACE, supporting independent expression during *R. prowazekii* infection of HMECs. Analysis of sRNA nucleotide sequences using computational approaches was employed to characterize the secondary structures and potential mRNAs targets of these novel sRNAs. The presence and expression of novel *R. prowazekii* sRNAs during human infection were established and potential functions suggested for these non-coding RNA regulators.

The third aim detailed the identification of novel *R. prowazekii* sRNAs expressed during the arthropod stage of its life cycle. In addition, it characterized and validated a predicted mRNA target for the confirmed *Rp*_sR60 sRNA expressed during *R.*

prowazekii infection of human microvascular endothelial cells. Next generation sequencing of AAE2 cells infected with *R. prowazekii* demonstrated the expression of 93 novel sRNAs, which 67 and 26 were considered *cis*-acting and *trans*-acting, respectively. Five novel sRNAs were selected, and their expression was compared between the arthropod vector and the human host using qRT-PCR. The analysis demonstrated that four of the five sRNAs (*Rp*_sR76, *Rp*_sR83, *Rp*_sR86, and *Rp*_sR159) were expressed at significantly higher levels at 3 h post-infection in AAE2 cells than HMECs. The fifth sRNA (*Rp*_sR63) showed no significant differences in expression at 3h or 24h post-infection. Previously, I validated the expression of *Rp*_sR60 during *R. prowazekii* infection of HMECs. A list of 53 potential mRNA target interactions was developed using the web-based CopraRNA. To validate *Rp*_sR60 interaction with the targets, I selected ten targets and analyzed their interactions using EMSA. From the list, only H375_0420 demonstrated a stable RNA complex at 30°C and 37°C, thereby confirming the gene as a target for *Rp*_sR60. Interestingly, during HMEC infection, RT-PCR presented a significant decrease in expression for both H375_0420 and *Rp*_sR60 after 0.5 h post-infection; however, there were no significant differences in expression during AAE2 infection. Differential expression of novel sRNAs between the arthropod vector and the human host was established, alluding to different mechanisms of bacterial maintenance during the *R. prowazekii* life cycle.

This project has provided a new understanding for rickettsial virulence, a subject for which little is known despite over a century of study. Due to reductive evolution, rickettsial gene expression and regulation remain rather unclear. Small RNAs allow bacteria to fine tune gene expression without having to invest energy or resources into

protein-based pathways. Therefore, sRNA elucidation in *Rickettsia* may lead to exciting new directions in rickettsial genetics. As over 150 sRNAs were found to be expressed during *R. prowazekii* infection of human microvascular endothelial cells and *Amblyomma* tick cells, their targets and the effects these sRNAs have on their targets remain unknown. Confirming and characterizing these roles will lead to a more complete picture of the virulence mechanisms employed by *Rickettsia prowazekii*, a pathogen that has had serious impacts on war, famine, homeless populations, and natural disasters (Raoult *et al.*, 2004; Bechah *et al.*, 2008a).

Future studies should also investigate and define the sRNA expression profile between different environmental conditions and niches. Particularly for spotted fever rickettsiae, an eschar, cutaneous necrosis that resulted from severe injury to the blood vessels, may form at the bite site of an infected arthropod (Parola *et al.*, 2005). An examination of *R. conorii* gene expression at an eschar site from infected human skin biopsies demonstrated that about 15% of *R. conorii* ORFs were differentially expressed (Renesto *et al.*, 2008). Of the 211 transcripts found to have differential expression when compared to *R. conorii* grown in Vero cells, 180 genes demonstrated a 2-fold or greater down-regulation, most of which were genes involved in translation, cell wall synthesis, secretion, and energy production. Conversely, 31 genes were up-regulated by two-fold or greater. These genes mostly belonged to DNA repair and modification and osmotic stress response. Interestingly, of these genes, 21 were absent from typhus group rickettsiae (Renesto *et al.*, 2008). Differences in gene expression were also noted in *R. rickettsii* grown at 37°C in Vero cells and 22°C in ISE6 cells (*Ixodes scapularis* [deer tick]). Using a 3-fold cut off, six rickettsial genes (3 up-regulated and 3 down-regulated) met or

exceeded this threshold, but are annotated as hypothetical proteins with no known homolog (Ellison *et al.*, 2009). When subjected to significant changes in prolonged temperatures (4°C for 2 h vs 4°C for 24 h), there were minimal differences. That is, the same gene expression profiles were found whether the *R. rickettsii* were at 4°C for 2 h or 24 h. However, when the gene expressions were compared between significant temperature changes with prolonged exposure (34°C vs 4°C), there were 56 genes that met or exceeded the 3-fold up- or down-regulated threshold (Ellison *et al.*, 2009). When shifting temperatures from 34°C to 42°C in order to heat shock *R. prowazekii,* there were 23 transcripts that met or exceeded at 2-fold threshold, which were all significantly increased. Of these, 57% were known heat-shock-inducible genes. There were no genes found that were reliably down-regulated during this study (Audia *et al.*, 2008). As a result, I hypothesize that interesting sRNA expression profiles will be observed under these regimens, providing a more thorough explanation to the above changes in gene expression. Further, significant differences will be observed between *R. prowazekii* sRNA profiles during tick cell infections and body louse infections.

Appendix Additional Files

Additional File 1. sRNA Predictions Using SIPHT

Rickettsia africae SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RAF_ORF0138	>>>	29	152429	152558	>>>	129	567	>>>	RAF_ORF0140
2	RAF_ORF0416	<<<	1	443156	443213	>>>	57	49	<<<	RAF_ORF0417
3	RAF_ORF0552	<<<	16	582515	582575	>>>	60	138	>>>	RAF_ORF0553
4	RAF_ORF0587	>>>	1	619636	619728	>>>	92	22	>>>	RAF_ORF0588
5	RAF_ORF0701	<<<	82	734968	735431	>>>	463	152	<<<	RAF_ORF0702
6	RAF_ORF1101	>>>	314	1115365	1115524	>>>	159	179	<<<	RAF_ORF1103
7	RAF_ORF1231	>>>	2	1257748	1257853	<<<	105	0	>>>	RAF_ORF1232
8	RAF_ORF0810	>>>	110	850274	850394	<<<	120	13	>>>	RAF_ORF0811
9	RAF_ORF0753	>>>	120	788757	789283	<<<	526	633	>>>	RAF_RNA22
10	RAF_ORF0706	>>>	105	739648	739741	<<<	93	176	<<<	RAF_ORF0707
11	RAF_ORF0622	<<<	594	662308	662399	<<<	91	244	>>>	RAF_ORF0623
12	RAF_ORF0546	>>>	512	577297	577640	<<<	343	194	<<<	RAF_ORF0547
13	RAF_ORF0325	<<<	1016	356820	356894	<<<	74	2	>>>	RAF_ORF0326
14	RAF_ORF0213	>>>	122	231244	231392	<<<	148	268	<<<	RAF_ORF0214
15	RAF_ORF0029	<<<	325	32306	32395	<<<	89	10	<<<	RAF_ORF0030
16	RAF_ORF0042	>>>	55	41898	42178	<<<	280	167	>>>	RAF_ORF0043
17	RAF_ORF0062	>>>	352	62195	62258	>>>	63	204	>>>	RAF_ORF0063
18	RAF_ORF0067	>>>	34	67255	67338	>>>	83	65	>>>	RAF_ORF0068
19	RAF_ORF0086	>>>	445	85846	85976	>>>	130	-10	<<<	RAF_ORF0087

20	RAF_ORF0093	<<<	2	93049	93320	<<<	271	78	<<<	RAF_ORF0094
21	RAF_ORF0098	>>>	19	97834	98058	<<<	224	0	<<<	RAF_ORF0099
22	RAF_ORF0104	>>>	76	104441	104968	<<<	527	297	<<<	RAF_ORF0105
23	RAF_ORF0107	>>>	69	113887	113952	>>>	65	61	>>>	RAF_ORF0108
24	RAF_ORF0109	>>>	63	115812	116214	>>>	402	245	<<<	RAF_ORF0110
25	RAF_ORF0121	>>>	222	127525	127879	>>>	354	-13	<<<	RAF_ORF0122
26	RAF_ORF0129	>>>	1	135204	135326	>>>	122	-7	>>>	RAF_ORF0130
27	RAF_ORF0155	>>>	2	167327	167463	>>>	136	37	>>>	RAF_ORF0156
28	RAF_ORF0160	>>>	1	172040	172180	>>>	140	18	>>>	RAF_ORF0161
29	RAF_ORF0165	>>>	1	176424	176594	>>>	170	99	>>>	RAF_ORF0166
30	RAF_ORF0169	>>>	675	179729	179783	>>>	54	214	>>>	RAF_ORF0170
31	RAF_ORF0170	>>>	16	184136	184195	>>>	59	181	>>>	RAF_ORF0171
32	RAF_ORF0171	>>>	340	188836	188901	<<<	65	0	<<<	RAF_ORF0172
33	RAF_ORF0171	>>>	1	188497	188896	>>>	399	5	<<<	RAF_ORF0172
34	RAF_ORF0176	>>>	141	195319	195359	<<<	40	95	<<<	RAF_ORF0177
35	RAF_ORF0213	>>>	297	231419	231475	<<<	56	185	<<<	RAF_ORF0214
36	RAF_ORF0213	>>>	148	231270	231459	>>>	189	201	<<<	RAF_ORF0214
37	RAF_ORF0218	<<<	57	235814	235920	<<<	106	0	<<<	RAF_ORF0219
38	RAF_ORF0218	<<<	78	235835	235920	<<<	85	0	<<<	RAF_ORF0219
39	RAF_ORF0220	<<<	40	240372	240492	<<<	120	15	<<<	RAF_ORF0221
40	RAF_ORF0226	>>>	862	249000	249051	<<<	51	245	<<<	RAF_ORF0227
41	RAF_ORF0226	>>>	1039	249177	249295	<<<	118	1	<<<	RAF_ORF0227
42	RAF_ORF0230	>>>	112	252068	252113	>>>	45	317	>>>	RAF_ORF0231
43	RAF_ORF0232	>>>	1	253871	253956	>>>	85	256	<<<	RAF_ORF0233
44	RAF_ORF0232	>>>	24	253894	254212	<<<	318	0	<<<	RAF_ORF0233
45	RAF_ORF0232	>>>	151	254021	254212	<<<	191	0	<<<	RAF_ORF0233

46	RAF_ORF0232	>>>	1	253871	254079	>>>	208	133	<<<	RAF_ORF0233
47	RAF_ORF0234	>>>	115	255867	255918	>>>	51	131	<<<	RAF_ORF0235
48	RAF_ORF0234	>>>	108	255860	255962	<<<	102	87	<<<	RAF_ORF0235
49	RAF_ORF0262	<<<	129	289639	289723	<<<	84	0	<<<	RAF_ORF0263
50	RAF_ORF0276	>>>	148	304650	304815	<<<	165	250	>>>	RAF_ORF0277
51	RAF_ORF0279	<<<	315	307664	307751	<<<	87	40	<<<	RAF_ORF0281
52	RAF_ORF0279	<<<	25	307374	307729	>>>	355	62	<<<	RAF_ORF0281
53	RAF_ORF0309	>>>	8	336644	336755	>>>	111	123	>>>	RAF_ORF0310
54	RAF_ORF0309	>>>	8	336644	336797	>>>	153	81	>>>	RAF_ORF0310
55	RAF_ORF0312	>>>	25	340033	340149	<<<	116	0	<<<	RAF_ORF0313
56	RAF_ORF0318	>>>	13	347222	347346	>>>	124	107	>>>	RAF_ORF0319
57	RAF_ORF0326	>>>	19	357471	357603	>>>	132	29	>>>	RAF_ORF0327
58	RAF_ORF0326	>>>	19	357471	357625	>>>	154	7	>>>	RAF_ORF0327
59	RAF_ORF0333	>>>	45	363633	363753	>>>	120	324	>>>	RAF_ORF0334
60	RAF_ORF0337	>>>	635	366763	366840	>>>	77	145	<<<	RAF_ORF0338
61	RAF_ORF0352	<<<	82	380465	380538	<<<	73	22	<<<	RAF_ORF0353
62	RAF_ORF0386	>>>	113	415355	415478	>>>	123	195	>>>	RAF_ORF0387
63	RAF_ORF0388	>>>	4	416157	416237	>>>	80	314	<<<	RAF_ORF0389
64	RAF_ORF0388	>>>	0	416153	416551	<<<	398	0	<<<	RAF_ORF0389
65	RAF_ORF0400	<<<	81	429567	429654	<<<	87	18	<<<	RAF_ORF0401
66	RAF_ORF0415	>>>	245	441699	441808	<<<	109	2	<<<	RAF_ORF0416
67	RAF_ORF0415	>>>	188	441642	441757	>>>	115	53	<<<	RAF_ORF0416
68	RAF_ORF0423	>>>	36	451086	451227	>>>	141	79	>>>	RAF_ORF0424
69	RAF_ORF0434	>>>	2291	465052	465261	<<<	209	0	<<<	RAF_ORF0435
70	RAF_ORF0461	>>>	6	488452	488571	>>>	119	60	<<<	RAF_ORF0462
71	RAF_ORF0461	>>>	6	488452	488553	>>>	101	78	<<<	RAF_ORF0462

72	RAF_ORF0464	>>>	154	492656	492712	>>>	56	240	>>>	RAF_ORF0465
73	RAF_ORF0473	>>>	1	500577	500675	>>>	98	39	>>>	RAF_ORF0474
74	RAF_ORF0482	>>>	678	511620	511769	>>>	149	302	<<<	RAF_ORF0483
75	RAF_ORF0500	<<<	89	531247	531383	<<<	136	6	<<<	RAF_ORF0501
76	RAF_ORF0508	>>>	1	537538	537676	>>>	138	77	>>>	RAF_ORF0509
77	RAF_ORF0519	>>>	11	551424	551506	>>>	82	99	>>>	RAF_ORF0520
78	RAF_ORF0519	>>>	11	551424	551550	>>>	126	55	>>>	RAF_ORF0520
79	RAF_ORF0525	>>>	751	558475	558576	>>>	101	114	<<<	RAF_ORF0526
80	RAF_ORF0525	>>>	788	558512	558611	<<<	99	79	<<<	RAF_ORF0526
81	RAF_ORF0525	>>>	751	558475	558556	>>>	81	134	<<<	RAF_ORF0526
82	RAF_ORF0544	>>>	199	573484	573574	>>>	90	70	>>>	RAF_ORF0545
83	RAF_ORF0546	>>>	849	577634	577697	<<<	63	137	<<<	RAF_ORF0547
84	RAF_ORF0550	<<<	31	580290	580403	<<<	113	0	<<<	RAF_ORF0551
85	RAF_ORF0555	>>>	1	585113	585241	>>>	128	2	>>>	RAF_ORF0556
86	RAF_ORF0561	<<<	67	589835	589921	<<<	86	0	<<<	RAF_ORF0563
87	RAF_ORF0570	>>>	513	596829	596875	<<<	46	37	<<<	RAF_ORF0571
88	RAF_ORF0570	>>>	471	596787	596899	>>>	112	13	<<<	RAF_ORF0571
89	RAF_ORF0575	>>>	3	602662	602826	>>>	164	33	>>>	RAF_ORF0576
90	RAF_ORF0578	>>>	23	606414	606459	>>>	45	426	>>>	RAF_ORF0579
91	RAF_ORF0582	>>>	1	612149	612286	>>>	137	37	<<<	RAF_ORF0583
92	RAF_ORF0584	>>>	4	615782	616090	>>>	308	1192	>>>	RAF_ORF0585
93	RAF_ORF0584	>>>	4	615782	616060	>>>	278	1222	>>>	RAF_ORF0585
94	RAF_ORF0587	>>>	1	619636	619751	>>>	115	-1	>>>	RAF_ORF0588
95	RAF_ORF0603	<<<	45	638738	638862	<<<	124	0	<<<	RAF_ORF0604
96	RAF_ORF0612	>>>	1	650817	650913	>>>	96	20	<<<	RAF_ORF0613
97	RAF_ORF0621	<<<	91	658915	659066	<<<	151	10	<<<	RAF_ORF0622

98	RAF_ORF0625	>>>	119	664914	664950	>>>	36	45	>>>	RAF_ORF0626
99	RAF_ORF0640	>>>	2	674913	675039	>>>	126	7	<<<	RAF_ORF0641
100	RAF_ORF0692	>>>	97	724559	724614	<<<	55	65	<<<	RAF_ORF0693
101	RAF_ORF0706	>>>	45	739588	739673	<<<	85	244	<<<	RAF_ORF0707
102	RAF_ORF0707	<<<	71	740532	740718	<<<	186	10	<<<	RAF_ORF0708
103	RAF_ORF0718	>>>	96	750272	750428	<<<	156	0	>>>	RAF_RNA21
104	RAF_ORF0718	>>>	1	750177	750332	>>>	155	96	>>>	RAF_RNA21
105	RAF_ORF0733	>>>	153	765520	765627	>>>	107	105	>>>	RAF_ORF0734
106	RAF_ORF0734	>>>	19	767237	767350	>>>	113	707	<<<	RAF_ORF0735
107	RAF_ORF0734	>>>	109	767327	767378	>>>	51	679	<<<	RAF_ORF0735
108	RAF_ORF0740	>>>	1	775232	775300	>>>	68	50	>>>	RAF_ORF0741
109	RAF_ORF0749	<<<	54	786395	786498	<<<	103	0	<<<	RAF_ORF0750
110	RAF_ORF0753	>>>	20	788657	788813	>>>	156	1103	>>>	RAF_RNA22
111	RAF_ORF0775	<<<	277	811665	811745	<<<	80	317	<<<	RAF_ORF0776
112	RAF_ORF0789	<<<	720	828111	828162	<<<	51	176	<<<	RAF_ORF0790
113	RAF_ORF0826	<<<	65	865438	865560	<<<	122	20	<<<	RAF_ORF0827
114	RAF_ORF0849	>>>	166	885817	886050	>>>	233	108	<<<	RAF_ORF0850
115	RAF_ORF0850	<<<	21	888121	888564	<<<	443	1129	<<<	RAF_ORF0851
116	RAF_ORF0850	<<<	220	888320	888497	<<<	177	1196	<<<	RAF_ORF0851
117	RAF_ORF0860	>>>	819	904832	905043	>>>	211	30	>>>	RAF_ORF0861
118	RAF_ORF0865	>>>	9	913242	913308	>>>	66	190	<<<	RAF_ORF0866
119	RAF_ORF0865	>>>	9	913242	913327	>>>	85	171	<<<	RAF_ORF0866
120	RAF_ORF0868	>>>	12	915483	915590	>>>	107	178	>>>	RAF_ORF0869
121	RAF_ORF0868	>>>	12	915483	915743	>>>	260	25	>>>	RAF_ORF0869
122	RAF_ORF0870	>>>	40	917884	917952	>>>	68	849	<<<	RAF_ORF0871
123	RAF_ORF0870	>>>	353	918197	918454	<<<	257	347	<<<	RAF_ORF0871

124	RAF_ORF0876	<<<	639	923936	923997	<<<	61	0	<<<	RAF_ORF0877
125	RAF_ORF0881	>>>	71	927982	928227	>>>	245	431	<<<	RAF_ORF0882
126	RAF_ORF0883	>>>	-26	930083	930303	<<<	220	20	<<<	RAF_RNA27
127	RAF_ORF0928	<<<	33	959307	959420	<<<	113	6	<<<	RAF_ORF0929
128	RAF_ORF0947	>>>	24	976827	976984	>>>	157	117	<<<	RAF_ORF0948
129	RAF_ORF0947	>>>	141	976944	977063	<<<	119	38	<<<	RAF_ORF0948
130	RAF_ORF0966	<<<	6	996979	997191	>>>	212	55	>>>	RAF_ORF0971
131	RAF_ORF0973	>>>	24	998936	999060	>>>	124	89	<<<	RAF_ORF0974
132	RAF_ORF0973	>>>	24	998936	999198	>>>	262	-49	<<<	RAF_ORF0974
133	RAF_ORF0978	>>>	7	1004563	1004694	>>>	131	163	>>>	RAF_ORF0979
134	RAF_ORF0979	>>>	586	1005693	1005754	>>>	61	18	>>>	RAF_ORF0980
135	RAF_ORF0986	<<<	337	1017019	1017087	<<<	68	0	<<<	RAF_ORF0990
136	RAF_ORF1011	>>>	142	1038188	1038230	<<<	42	169	<<<	RAF_ORF1012
137	RAF_ORF1020	>>>	1	1043691	1043818	>>>	127	-3	>>>	RAF_ORF1021
138	RAF_RNA33	>>>	1	1050625	1050975	>>>	350	22	<<<	RAF_ORF1026
139	RAF_ORF1026	<<<	1231	1052595	1052862	<<<	267	1776	<<<	RAF_ORF1027
140	RAF_ORF1037	>>>	57	1062192	1062317	<<<	125	339	<<<	RAF_ORF1038
141	RAF_ORF1048	<<<	65	1070883	1070937	>>>	54	525	>>>	RAF_ORF1050
142	RAF_ORF1080	>>>	123	1096976	1097063	>>>	87	70	>>>	RAF_ORF1081
143	RAF_ORF1082	>>>	50	1100113	1100327	<<<	214	0	<<<	RAF_ORF1083
144	RAF_ORF1082	>>>	1	1100064	1100369	>>>	305	-42	<<<	RAF_ORF1083
145	RAF_ORF1087	<<<	21	1102677	1102895	<<<	218	14	<<<	RAF_ORF1088
146	RAF_ORF1119	>>>	2	1143312	1143601	>>>	289	186	<<<	RAF_ORF1120
147	RAF_ORF1121	<<<	57	1145531	1145594	<<<	63	0	<<<	RAF_ORF1122
148	RAF_ORF1144	>>>	1	1169670	1169870	>>>	200	-57	<<<	RAF_ORF1145
149	RAF_ORF1144	>>>	1	1169670	1169827	>>>	157	-14	<<<	RAF_ORF1145

150	RAF_ORF1173	>>>	6	1198512	1198879	>>>	367	13	<<<	RAF_ORF1174
151	RAF_ORF1173	>>>	321	1198827	1198892	<<<	65	0	<<<	RAF_ORF1174
152	RAF_ORF1193	>>>	54	1217345	1217472	>>>	127	1472	>>>	RAF_ORF1194
153	RAF_ORF1208	<<<	137	1235008	1235084	<<<	76	1	<<<	RAF_ORF1211
154	RAF_ORF1192	>>>	137	1216005	1216045	<<<	40	0	>>>	RAF_ORF1193
155	RAF_ORF0118	>>>	35	125254	125349	<<<	95	14	>>>	RAF_ORF0119
156	RAF_ORF0129	>>>	1	135204	135296	>>>	92	23	>>>	RAF_ORF0130
157	RAF_ORF0177	<<<	245	197815	197950	<<<	135	2	<<<	RAF_ORF0178
158	RAF_ORF0206	>>>	212	226782	226832	>>>	50	100	<<<	RAF_ORF0207
159	RAF_ORF0226	>>>	221	248359	248554	<<<	195	742	<<<	RAF_ORF0227
160	RAF_ORF0333	>>>	45	363633	363716	>>>	83	361	>>>	RAF_ORF0334
161	RAF_ORF0489	>>>	87	519562	519605	<<<	43	33	>>>	RAF_ORF0490
162	RAF_RNA17	>>>	-38	699363	699616	<<<	253	484	>>>	RAF_ORF0665
163	RAF_ORF0746	<<<	16	783642	783748	<<<	106	998	<<<	RAF_ORF0747

Rickettsia akari SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	A1C_00730	>>>	10	131924	131980	>>>	56	142	>>>	A1C_00735
2	A1C_00760	>>>	120	136143	136676	>>>	533	233	>>>	A1C_00780
3	A1C_01315	<<<	1	248816	249066	>>>	250	15	<<<	A1C_01320
5	A1C_03440	>>>	1	621612	621786	>>>	174	22	>>>	A1C_03445
6	A1C_05510	>>>	105	977703	977924	>>>	221	-14	>>>	A1C_05515
7	A1C_05515	>>>	105	979675	979972	>>>	297	205	<<<	A1C_05530
9	A1C_05945	>>>	19	1056399	1056583	>>>	184	1243	>>>	A1C_05960
10	A1C_06710	<<<	4	1207113	1207187	>>>	74	154	<<<	A1C_06715
11	A1C_06865	<<<	739	1227903	1228122	<<<	219	59	<<<	A1C_06870
12	A1C_06125	>>>	146	1101212	1101245	<<<	33	68	>>>	A1C_06130
13	A1C_05620	>>>	314	999762	999925	<<<	163	23	<<<	A1C_05625
14	A1C_04715	<<<	292	853273	853541	<<<	268	471	<<<	A1C_04730
15	A1C_04440	<<<	963	802244	802295	<<<	51	6	>>>	A1C_04455
16	A1C_03640	<<<	491	661673	661761	<<<	88	266	>>>	A1C_03645
17	A1C_02080	>>>	58	391653	391790	<<<	137	1585	>>>	A1C_02105
18	A1C_01640	>>>	119	308494	308593	<<<	99	168	>>>	A1C_01645
19	A1C_00710	<<<	91	129928	130176	<<<	248	0	>>>	A1C_00715
20	A1C_00420	<<<	211	67071	67115	<<<	44	253	>>>	A1C_00425
21	A1C_00185	>>>	57	26255	26318	<<<	63	21	>>>	A1C_00190
22	A1C_00115	>>>	428	18473	18678	<<<	205	31	<<<	A1C_00120
23	A1C_00115	>>>	478	18523	18655	>>>	132	54	<<<	A1C_00120
24	A1C_00115	>>>	478	18523	18625	>>>	102	84	<<<	A1C_00120
25	A1C_00155	<<<	365	21995	22066	<<<	71	48	<<<	A1C_00160
26	A1C_00320	<<<	208	49724	49798	<<<	74	203	>>>	A1C_t06889
27	A1C_00365	>>>	1	57409	57545	>>>	136	439	>>>	A1C_00370
28	A1C_00430	>>>	53	67987	68125	>>>	138	246	<<<	A1C_00435
29	A1C_00430	>>>	53	67987	68107	>>>	120	264	<<<	A1C_00435
30	A1C_00530	<<<	36	88654	88896	<<<	242	253	<<<	A1C_00535
31	A1C_00530	<<<	291	88909	89147	<<<	238	2	<<<	A1C_00535

32	A1C_00530	<<<	383	89001	89147	<<<	146	2	<<<	A1C_00535
33	A1C_00605	<<<	74	106739	107060	<<<	321	85	<<<	A1C_00610
34	A1C_00705	<<<	173	128504	128604	<<<	100	38	<<<	A1C_00710
35	A1C_00730	>>>	10	131924	132001	>>>	77	121	>>>	A1C_00735
36	A1C_00780	>>>	15	137213	137466	>>>	253	29	>>>	A1C_00785
37	A1C_01005	>>>	134	185054	185216	>>>	162	587	>>>	A1C_01010
38	A1C_01015	>>>	300	194663	194792	>>>	129	-4	<<<	A1C_01020
39	A1C_01315	<<<	149	248964	249051	<<<	87	30	<<<	A1C_01320
40	A1C_01335	>>>	143	254567	254639	>>>	72	121	>>>	A1C_01340
41	A1C_01340	>>>	132	255385	255507	<<<	122	0	<<<	A1C_01345
42	A1C_01375	>>>	26	260290	260560	<<<	270	1	<<<	A1C_01380
44	A1C_01450	>>>	1	269536	269649	>>>	113	59	<<<	A1C_01455
45	A1C_01525	>>>	233	283481	283553	>>>	72	229	<<<	A1C_01530
46	A1C_01605	>>>	46	298516	298557	>>>	41	121	>>>	A1C_01610
47	A1C_01815	>>>	5	339807	339917	>>>	110	32	>>>	A1C_01820
48	A1C_01815	>>>	5	339807	339893	>>>	86	56	>>>	A1C_01820
49	A1C_01965	>>>	702	368575	368809	<<<	234	85	<<<	A1C_01970
50	A1C_01965	>>>	689	368562	368627	>>>	65	267	<<<	A1C_01970
51	A1C_02470	>>>	20	451678	451838	>>>	160	33	>>>	A1C_02475
52	A1C_02555	<<<	288	465579	465720	<<<	141	39	<<<	A1C_02560
53	A1C_02595	>>>	1	470159	470226	>>>	67	218	<<<	A1C_02600
54	A1C_02630	>>>	125	477844	477994	>>>	150	119	<<<	A1C_02635
55	A1C_02795	>>>	3	508289	508450	>>>	161	333	<<<	A1C_02800
56	A1C_02795	>>>	3	508289	508420	>>>	131	363	<<<	A1C_02800
57	A1C_02825	<<<	1040	513630	513721	<<<	91	108	<<<	A1C_02830
58	A1C_02935	<<<	142	531589	531855	<<<	266	0	<<<	A1C_02940
59	A1C_03035	>>>	95	552768	552890	>>>	122	96	>>>	A1C_03040
60	A1C_03035	>>>	95	552768	552872	>>>	104	114	>>>	A1C_03040
61	A1C_03065	>>>	5	558790	558892	>>>	102	208	>>>	A1C_03070
62	A1C_03070	>>>	16	559330	559405	>>>	75	2168	<<<	A1C_03095
63	A1C_03070	>>>	1916	561230	561573	<<<	343	0	<<<	A1C_03095

64	A1C_03180	<<<	126	579805	579908	<<<	103	2	<<<	A1C_03185
65	A1C_03210	>>>	189	585613	585709	<<<	96	91	<<<	A1C_03215
66	A1C_03365	>>>	78	606913	607053	>>>	140	78	>>>	A1C_03370
67	A1C_03365	>>>	78	606913	607026	>>>	113	105	>>>	A1C_03370
68	A1C_03420	<<<	91	618989	619321	<<<	332	1071	<<<	A1C_03435
69	A1C_03440	>>>	1	621612	621815	>>>	203	-7	>>>	A1C_03445
70	A1C_03610	>>>	42	654084	654206	>>>	122	254	>>>	A1C_03615
71	A1C_03875	>>>	127	696545	696597	<<<	52	99	<<<	A1C_03880
72	A1C_04005	>>>	16	720009	720122	>>>	113	56	>>>	A1C_04010
73	A1C_t06917	>>>	116	754294	754349	>>>	55	133	>>>	A1C_04190
74	A1C_04240	<<<	135	762017	762110	<<<	93	32	<<<	A1C_04245
75	A1C_04270	>>>	70	766086	766238	>>>	152	435	<<<	A1C_04275
76	A1C_04365	>>>	25	786350	786477	>>>	127	102	<<<	A1C_04370
77	A1C_04365	>>>	25	786350	786459	>>>	109	120	<<<	A1C_04370
78	A1C_04490	>>>	9	809378	809538	>>>	160	86	<<<	A1C_04495
79	A1C_04505	<<<	102	811921	811988	<<<	67	169	<<<	A1C_04510
80	A1C_04520	>>>	178	817885	817970	<<<	85	101	<<<	A1C_04525
81	A1C_04575	>>>	1	824967	825071	>>>	104	21	>>>	A1C_04580
82	A1C_04580	>>>	15	825630	825717	>>>	87	129	>>>	A1C_04585
83	A1C_04595	>>>	695	829469	829546	<<<	77	80	<<<	A1C_04600
84	A1C_04745	<<<	127	856379	856483	<<<	104	0	<<<	A1C_04750
85	A1C_04780	>>>	48	860015	860105	>>>	90	308	<<<	A1C_04785
86	A1C_04815	<<<	323	867017	867114	<<<	97	8	<<<	A1C_04820
87	A1C_04900	>>>	127	887289	887389	<<<	100	450	<<<	A1C_04905
88	A1C_05300	>>>	97	944498	944586	>>>	88	397	<<<	A1C_05305
89	A1C_05300	>>>	174	944575	944791	<<<	216	192	<<<	A1C_05305
90	A1C_05300	>>>	318	944719	944791	<<<	72	192	<<<	A1C_05305
91	A1C_05300	>>>	356	944757	944791	<<<	34	192	<<<	A1C_05305
92	A1C_05300	>>>	486	944887	944966	<<<	79	17	<<<	A1C_05305
93	A1C_05300	>>>	24	944425	944927	>>>	502	56	<<<	A1C_05305
94	A1C_05475	>>>	9	969868	969979	>>>	111	207	>>>	A1C_05480

95	A1C_05620	>>>	231	999679	999850	>>>	171	98	<<<	A1C_05625
96	A1C_05820	>>>	124	1033065	1033142	<<<	77	64	<<<	A1C_05825
97	A1C_05900	>>>	239	1048905	1049127	<<<	222	19	<<<	A1C_05905
98	A1C_05910	<<<	1099	1050726	1050830	<<<	104	386	<<<	A1C_05915
99	A1C_05945	>>>	428	1056808	1057253	>>>	445	573	>>>	A1C_05960
100	A1C_06035	>>>	33	1072075	1072236	>>>	161	418	>>>	A1C_06040
101	A1C_06185	<<<	70	1111199	1111315	<<<	116	1	<<<	A1C_06190
102	A1C_06295	>>>	205	1130080	1130464	>>>	384	167	>>>	A1C_t06945
103	A1C_06340	<<<	109	1140006	1140266	<<<	260	735	<<<	A1C_06355
104	A1C_06395	>>>	31	1147816	1147906	>>>	90	980	>>>	A1C_06410
105	A1C_06395	>>>	394	1148179	1148336	>>>	157	550	>>>	A1C_06410
106	A1C_06415	>>>	168	1150362	1150490	>>>	128	138	<<<	A1C_06420
108	A1C_06605	<<<	191	1186343	1186441	<<<	98	77	<<<	A1C_06610
109	A1C_06680	>>>	7	1201461	1201550	>>>	89	69	<<<	A1C_06685
110	A1C_06705	>>>	92	1206251	1206311	<<<	60	101	<<<	A1C_06710
111	A1C_06860	>>>	99	1226827	1226918	>>>	91	125	<<<	A1C_06865
112	A1C_06860	>>>	99	1226827	1227059	>>>	232	-16	<<<	A1C_06865
113	A1C_05945	>>>	318	1056698	1056791	<<<	93	1035	>>>	A1C_05960
114	A1C_06080	>>>	39	1089027	1089171	<<<	144	71	>>>	A1C_06085
115	A1C_06415	>>>	365	1150559	1150610	<<<	51	18	<<<	A1C_06420
116	A1C_06415	>>>	168	1150362	1150594	>>>	232	34	<<<	A1C_06420
117	A1C_01495	>>>	-4	278105	278173	<<<	68	2377	>>>	A1C_01525
118	A1C_04075	<<<	16	737330	737436	<<<	106	216	<<<	A1C_04080
119	A1C_05365	>>>	635	957860	958189	<<<	329	1306	>>>	A1C_05400
120	A1C_05530	<<<	775	981070	981136	<<<	66	0	<<<	A1C_05535

Rickettsia bellii strain OSU

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	A1I_01875	>>>	206	364852	365124	>>>	272	11	>>>	A1I_01880
2	A1I_02655	<<<	6	492718	492823	>>>	105	46	<<<	A1I_02660
3	A1I_06135	<<<	64	1176398	1176541	>>>	143	119	<<<	A1I_06140
4	A1I_07920	<<<	25	1523781	1523864	>>>	83	39	<<<	A1I_07925
5	A1I_r08025	<<<	86	1126283	1126472	<<<	189	0	<<<	A1I_r01
6	A1I_01860	>>>	882	362976	363109	<<<	133	1032	>>>	A1I_01875
7	A1I_01250	<<<	290	247963	248045	<<<	82	309	>>>	A1I_01255
8	A1I_00070	>>>	78	15057	15160	<<<	103	63	<<<	A1I_00075
9	A1I_00070	>>>	4	14983	15101	>>>	118	122	<<<	A1I_00075
10	A1I_00370	>>>	81	67246	67347	<<<	101	97	<<<	A1I_00375
11	A1I_00395	>>>	345	72492	72553	<<<	61	114	<<<	A1I_00400
12	A1I_t07959	>>>	64	87346	87454	>>>	108	84	>>>	A1I_00480
13	A1I_t07959	>>>	64	87346	87433	>>>	87	105	>>>	A1I_00480
14	A1I_00910	<<<	274	167396	167459	<<<	63	36	<<<	A1I_00915
15	A1I_00950	>>>	41	175620	175709	>>>	89	62	>>>	A1I_00955
16	A1I_01030	>>>	236	203048	203167	>>>	119	1218	<<<	A1I_01045
17	A1I_01030	>>>	290	203102	203205	<<<	103	1180	<<<	A1I_01045
18	A1I_01030	>>>	236	203048	203146	>>>	98	1239	<<<	A1I_01045
19	A1I_01185	<<<	235	229219	229292	<<<	73	60	<<<	A1I_01190
20	A1I_01235	>>>	3167	244384	244519	>>>	135	516	<<<	A1I_01250
21	A1I_01290	>>>	3	255135	255208	>>>	73	349	<<<	A1I_01295
22	A1I_01290	>>>	127	255259	255319	>>>	60	238	<<<	A1I_01295
23	A1I_01290	>>>	187	255319	255351	>>>	32	206	<<<	A1I_01295
24	A1I_01730	>>>	67	335395	335565	>>>	170	418	>>>	A1I_t07965

25	A1I_01815	>>>	3	350140	350306	>>>	166	-35	>>>	A1I_01820
26	A1I_01875	>>>	206	364852	365141	>>>	289	-6	>>>	A1I_01880
27	A1I_01880	>>>	37	365353	365484	>>>	131	103	>>>	A1I_01885
28	A1I_01955	>>>	172	376512	376813	<<<	301	11	<<<	A1I_01960
29	A1I_01955	>>>	201	376541	376583	>>>	42	241	<<<	A1I_01960
30	A1I_02000	<<<	281	383406	383457	<<<	51	78	<<<	A1I_02005
31	A1I_02200	<<<	2654	410855	411210	<<<	355	979	<<<	A1I_02215
32	A1I_02560	>>>	490	477974	478114	>>>	140	116	<<<	A1I_02565
33	A1I_02560	>>>	566	478050	478148	<<<	98	82	<<<	A1I_02565
34	A1I_02560	>>>	490	477974	478090	>>>	116	140	<<<	A1I_02565
35	A1I_02630	>>>	50	488183	488348	>>>	165	478	<<<	A1I_02635
36	A1I_02630	>>>	50	488183	488318	>>>	135	508	<<<	A1I_02635
37	A1I_02630	>>>	141	488274	488462	<<<	188	364	<<<	A1I_02635
38	A1I_02630	>>>	283	488416	488462	<<<	46	364	<<<	A1I_02635
39	A1I_02710	>>>	282	504967	505101	<<<	134	290	<<<	A1I_02715
40	A1I_02970	>>>	51	552730	552771	>>>	41	793	>>>	A1I_02975
41	A1I_03130	>>>	174	581506	581539	>>>	33	1607	<<<	A1I_03145
42	A1I_03130	>>>	197	581529	581672	>>>	143	1474	<<<	A1I_03145
43	A1I_03130	>>>	284	581616	582058	<<<	442	1088	<<<	A1I_03145
44	A1I_03130	>>>	536	581868	582058	<<<	190	1088	<<<	A1I_03145
45	A1I_03225	>>>	127	597240	597632	<<<	392	66	<<<	A1I_03230
46	A1I_03360	>>>	45	623652	623785	>>>	133	90	<<<	A1I_03365
47	A1I_03360	>>>	134	623741	623827	<<<	86	48	<<<	A1I_03365
48	A1I_03490	>>>	1	648760	648880	>>>	120	106	>>>	A1I_03495
49	A1I_03495	>>>	117	651591	651696	>>>	105	134	>>>	A1I_03500
50	A1I_03620	>>>	1	675774	676075	>>>	301	2349	>>>	A1I_t07981

51	A1I_03685	>>>	53	691881	692051	>>>	170	147	>>>	A1I_03690
52	A1I_03685	>>>	53	691881	692030	>>>	149	168	>>>	A1I_03690
53	A1I_03700	<<<	254	693728	693782	<<<	54	55	<<<	A1I_03705
54	A1I_03780	>>>	181	711112	711214	>>>	102	101	<<<	A1I_03785
55	A1I_03780	>>>	181	711112	711193	>>>	81	122	<<<	A1I_03785
56	A1I_03925	>>>	12	738224	738335	>>>	111	94	>>>	A1I_03930
57	A1I_04080	>>>	112	767777	767822	>>>	45	291	>>>	A1I_04085
58	A1I_04080	>>>	112	767777	767877	>>>	100	236	>>>	A1I_04085
59	A1I_04080	>>>	112	767777	767847	>>>	70	266	>>>	A1I_04085
60	A1I_04270	<<<	18	805795	805905	<<<	110	48	<<<	A1I_04275
61	A1I_04395	>>>	38	833938	834091	>>>	153	103	>>>	A1I_04400
62	A1I_04395	>>>	38	833938	834070	>>>	132	124	>>>	A1I_04400
63	A1I_04550	<<<	223	860329	860437	<<<	108	20	<<<	A1I_04555
64	A1I_04980	>>>	30	943952	944059	>>>	107	112	>>>	A1I_04985
65	A1I_05425	>>>	24	1034748	1034857	>>>	109	277	<<<	A1I_05430
66	A1I_05425	>>>	277	1035001	1035135	>>>	134	-1	<<<	A1I_05430
67	A1I_05435	>>>	241	1036273	1036377	<<<	104	54	<<<	A1I_05440
68	A1I_05675	>>>	99	1089416	1089527	>>>	111	142	>>>	A1I_05680
69	A1I_05675	>>>	92	1089409	1089626	>>>	217	43	>>>	A1I_05680
70	A1I_05770	>>>	11	1105268	1105413	>>>	145	22	>>>	A1I_05775
71	A1I_05770	>>>	11	1105268	1105384	>>>	116	51	>>>	A1I_05775
72	A1I_06035	>>>	93	1156002	1156142	>>>	140	170	<<<	A1I_06040
73	A1I_06035	>>>	168	1156077	1156210	<<<	133	102	<<<	A1I_06040
74	A1I_06035	>>>	93	1156002	1156121	>>>	119	191	<<<	A1I_06040
75	A1I_06050	>>>	23	1160249	1160592	<<<	343	89	<<<	A1I_06055
76	A1I_06050	>>>	21	1160247	1160307	>>>	60	374	<<<	A1I_06055

77	A1I_06050	>>>	178	1160404	1160524	>>>	120	157	<<<	A1I_06055
78	A1I_06050	>>>	233	1160459	1160592	<<<	133	89	<<<	A1I_06055
79	A1I_06050	>>>	178	1160404	1160503	>>>	99	178	<<<	A1I_06055
80	A1I_06385	<<<	281	1221884	1221935	<<<	51	272	<<<	A1I_06390
81	A1I_t08011	<<<	121	1261906	1262072	>>>	166	117	<<<	A1I_06600
82	A1I_t08011	<<<	121	1261906	1262051	>>>	145	138	<<<	A1I_06600
83	A1I_06695	>>>	315	1283297	1283567	>>>	270	13	>>>	A1I_06700
84	A1I_06750	<<<	136	1293432	1293538	<<<	106	650	<<<	A1I_06755
85	A1I_06925	>>>	1347	1327097	1327248	<<<	151	1051	<<<	A1I_06940
86	A1I_06925	>>>	2380	1328130	1328252	>>>	122	47	<<<	A1I_06940
87	A1I_07055	>>>	4	1354527	1354625	>>>	98	161	>>>	A1I_07060
88	A1I_07055	>>>	4	1354527	1354604	>>>	77	182	>>>	A1I_07060
89	A1I_07115	>>>	43	1368479	1368594	>>>	115	98	>>>	A1I_07120
90	A1I_07140	>>>	32	1372455	1372529	>>>	74	6370	>>>	A1I_07160
91	A1I_07140	>>>	32	1372455	1372502	>>>	47	6397	>>>	A1I_07160
92	A1I_07200	<<<	36	1385852	1386113	<<<	261	428	<<<	A1I_07205
93	A1I_07250	<<<	173	1396624	1396819	<<<	195	397	<<<	A1I_07255
94	A1I_07270	<<<	188	1400556	1400657	<<<	101	74	<<<	A1I_t08017
95	A1I_t08017	<<<	259	1401068	1401133	<<<	65	1	<<<	A1I_07275
96	A1I_07360	>>>	51	1414637	1414723	>>>	86	315	>>>	A1I_07365
97	A1I_07270	<<<	81	1400449	1400657	<<<	208	74	<<<	A1I_t08017
98	A1I_01460	>>>	155	283482	283523	<<<	41	81	>>>	A1I_01465
99	A1I_03620	>>>	2171	677944	678091	<<<	147	333	>>>	A1I_t07981
100	A1I_03925	>>>	26	738238	738391	<<<	153	38	>>>	A1I_03930

Rickettsia bellii strain RML

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RBE_0347	<<<	81	392587	392864	>>>	277	24	>>>	RBE_0348
2	RBE_RNA08	>>>	22	420992	421180	>>>	188	86	>>>	RBE_RNA09
3	RBE_0446	<<<	59	505298	505595	>>>	297	24	>>>	RBE_0447
4	RBE_1204	<<<	311	1294260	1294340	>>>	80	290	>>>	RBE_1205
5	RBE_1423	<<<	25	1516877	1516960	>>>	83	39	<<<	RBE_1424
6	RBE_0852	>>>	46	940008	940115	<<<	107	4	>>>	RBE_0853
7	RBE_0671	<<<	824	751535	751701	<<<	166	29	>>>	RBE_0672
8	RBE_0329	>>>	120	370644	370789	<<<	145	143	>>>	RBE_0330
9	RBE_0015	>>>	78	15004	15107	<<<	103	63	<<<	RBE_0016
10	RBE_0015	>>>	4	14930	15048	>>>	118	122	<<<	RBE_0016
11	RBE_0088	>>>	1	101613	102010	>>>	397	81	>>>	RBE_0089
12	RBE_0097	>>>	142	110289	110414	>>>	125	76	>>>	RBE_0098
13	RBE_0097	>>>	142	110289	110393	>>>	104	97	>>>	RBE_0098
14	RBE_0113	<<<	331	133493	133565	<<<	72	49	<<<	RBE_0114
15	RBE_0119	>>>	221	138625	139073	>>>	448	-1	<<<	RBE_0120
16	RBE_RNA03	>>>	76	147543	147679	>>>	136	151	>>>	RBE_0130
17	RBE_RNA03	>>>	76	147543	147658	>>>	115	172	>>>	RBE_0130
18	RBE_RNA03	>>>	76	147543	147794	>>>	251	36	>>>	RBE_0130
19	RBE_0132	>>>	399	151381	151590	>>>	209	142	>>>	RBE_0133
20	RBE_0141	>>>	430	162093	162375	>>>	282	13	>>>	RBE_0142
21	RBE_0153	<<<	114	179587	179688	<<<	101	41	<<<	RBE_0154
22	RBE_0186	>>>	61	219889	219996	<<<	107	51	<<<	RBE_0187
23	RBE_RNA06	<<<	132	266213	266367	>>>	154	117	<<<	RBE_0223
24	RBE_RNA06	<<<	132	266213	266346	>>>	133	138	<<<	RBE_0223

25	RBE_0241	>>>	315	287468	287738	>>>	270	13	>>>	RBE_0242
26	RBE_0270	>>>	154	316092	316234	>>>	142	255	>>>	RBE_0271
27	RBE_0282	>>>	306	325911	325965	>>>	54	247	>>>	RBE_0283
28	RBE_0282	>>>	306	325911	325944	>>>	33	268	>>>	RBE_0283
29	RBE_0310	>>>	18	348598	348694	>>>	96	59	>>>	RBE_0311
30	RBE_0310	>>>	18	348598	348673	>>>	75	80	>>>	RBE_0311
31	RBE_0339	>>>	122	380199	380253	>>>	54	247	>>>	RBE_0340
32	RBE_0339	>>>	122	380199	380232	>>>	33	268	>>>	RBE_0340
33	RBE_0344	>>>	91	386707	386854	>>>	147	217	<<<	RBE_0345
34	RBE_0344	>>>	194	386810	386895	<<<	85	176	<<<	RBE_0345
35	RBE_0344	>>>	91	386707	387071	>>>	364	0	<<<	RBE_0345
36	RBE_0346	>>>	104	391201	391348	>>>	147	152	<<<	RBE_0347
37	RBE_0346	>>>	207	391304	391409	<<<	105	91	<<<	RBE_0347
38	RBE_0397	<<<	67	442809	442909	<<<	100	11	<<<	RBE_0398
39	RBE_0412	<<<	59	458697	458893	<<<	196	97	<<<	RBE_0413
40	RBE_0412	<<<	158	458796	458893	<<<	97	97	<<<	RBE_0413
41	RBE_0452	>>>	56	512216	512353	>>>	137	206	<<<	RBE_0453
42	RBE_0452	>>>	56	512216	512332	>>>	116	227	<<<	RBE_0453
43	RBE_0454	>>>	50	513507	513543	>>>	36	324	<<<	RBE_0455
44	RBE_0454	>>>	24	513481	513592	<<<	111	275	<<<	RBE_0455
45	RBE_0454	>>>	305	513762	513845	<<<	83	22	<<<	RBE_0455
46	RBE_0456	>>>	42	515430	515483	>>>	53	216	>>>	RBE_0457
47	RBE_0456	>>>	42	515430	515544	>>>	114	155	>>>	RBE_0457
48	RBE_RNA14	>>>	41	539345	539624	>>>	279	325	>>>	RBE_0473
49	RBE_RNA14	>>>	344	539648	539681	>>>	33	268	>>>	RBE_0473
50	RBE_0517	>>>	94	578765	578806	>>>	41	793	>>>	RBE_0518

51	RBE_0517	>>>	94	578765	578861	>>>	96	738	>>>	RBE_0518
52	RBE_0521	>>>	60	584241	584551	<<<	310	3	<<<	RBE_0522
53	RBE_0521	>>>	174	584355	584458	>>>	103	96	<<<	RBE_0522
54	RBE_0521	>>>	233	584414	584532	<<<	118	22	<<<	RBE_0522
55	RBE_0572	>>>	7	631471	631566	>>>	95	756	>>>	RBE_0573
56	RBE_0589	>>>	134	651382	651468	<<<	86	48	<<<	RBE_0590
57	RBE_0589	>>>	45	651293	651426	>>>	133	90	<<<	RBE_0590
58	RBE_0612	>>>	1	676412	676532	>>>	120	106	>>>	RBE_0613
59	RBE_0613	>>>	117	679243	679348	>>>	105	128	>>>	RBE_0614
60	RBE_0619	<<<	426	687110	687534	<<<	424	160	<<<	RBE_0620
61	RBE_0630	<<<	138	696279	696399	<<<	120	36	<<<	RBE_0631
62	RBE_0653	>>>	50	725208	725347	>>>	139	-13	>>>	RBE_0654
63	RBE_0773	>>>	30	855489	855596	>>>	107	112	>>>	RBE_0774
64	RBE_0823	>>>	603	910896	910953	>>>	57	107	>>>	RBE_0824
65	RBE_0823	>>>	603	910896	910932	>>>	36	128	>>>	RBE_0824
66	RBE_0841	>>>	292	927723	927871	>>>	148	266	<<<	RBE_0842
67	RBE_0855	>>>	126	944368	944447	>>>	79	391	<<<	RBE_0856
68	RBE_0855	>>>	126	944368	944426	>>>	58	412	<<<	RBE_0856
69	RBE_0855	>>>	126	944368	944591	>>>	223	247	<<<	RBE_0856
70	RBE_0855	>>>	126	944368	944570	>>>	202	268	<<<	RBE_0856
71	RBE_0855	>>>	284	944526	944647	<<<	121	191	<<<	RBE_0856
72	RBE_0861	<<<	193	951463	951694	<<<	231	49	<<<	RBE_0862
73	RBE_0867	>>>	84	954609	954740	>>>	131	530	<<<	RBE_0868
74	RBE_0867	>>>	154	954679	954782	<<<	103	488	<<<	RBE_0868
75	RBE_0867	>>>	84	954609	954719	>>>	110	551	<<<	RBE_0868
76	RBE_0889	>>>	63	978670	978767	>>>	97	246	>>>	RBE_0890

77	RBE_RNA22	>>>	106	992604	992775	>>>	171	727	>>>	RBE_0902
78	RBE_0947	<<<	110	1042307	1042404	<<<	97	9	<<<	RBE_0948
79	RBE_0974	>>>	139	1069452	1069518	<<<	66	179	<<<	RBE_0975
80	RBE_0993	<<<	184	1088435	1088570	<<<	135	51	<<<	RBE_0994
81	RBE_1068	>>>	80	1158844	1158930	>>>	86	439	>>>	RBE_1069
82	RBE_1068	>>>	80	1158844	1158909	>>>	65	460	>>>	RBE_1069
83	RBE_1075	<<<	427	1165735	1165784	<<<	49	118	<<<	RBE_1076
84	RBE_1090	<<<	266	1176688	1176790	<<<	102	703	<<<	RBE_1091
85	RBE_1090	<<<	602	1177024	1177283	<<<	259	210	<<<	RBE_1091
86	RBE_RNA33	<<<	446	1206603	1206747	<<<	144	65	<<<	RBE_1117
87	RBE_1197	>>>	230	1286971	1287041	<<<	70	125	<<<	RBE_1198
88	RBE_1197	>>>	267	1287008	1287082	<<<	74	84	<<<	RBE_1198
89	RBE_1197	>>>	361	1287102	1287165	<<<	63	1	<<<	RBE_1198
90	RBE_1224	<<<	235	1316708	1316781	<<<	73	60	<<<	RBE_1225
91	RBE_1242	<<<	274	1333099	1333162	<<<	63	36	<<<	RBE_1243
92	RBE_1265	<<<	75	1366110	1366188	<<<	78	39	<<<	RBE_1266
93	RBE_1277	<<<	119	1378893	1378968	<<<	75	62	<<<	RBE_RNA37
94	RBE_1293	>>>	147	1393821	1393886	>>>	65	308	<<<	RBE_1294
95	RBE_1293	>>>	147	1393821	1394058	<<<	237	136	<<<	RBE_1294
96	RBE_1293	>>>	147	1393821	1393865	>>>	44	329	<<<	RBE_1294
97	RBE_1298	>>>	99	1398996	1399111	>>>	115	65	<<<	RBE_1299
98	RBE_0947	<<<	40	1042237	1042389	>>>	152	24	<<<	RBE_0948
99	RBE_RNA03	>>>	76	147543	147750	>>>	207	80	>>>	RBE_0130
100	RBE_0400	<<<	10	445162	445317	>>>	155	151	>>>	RBE_0401

Rickettsia canadensis SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	A1E_00490	>>>	20	111277	111501	>>>	224	235	>>>	A1E_00495
2	A1E_00665	>>>	1	152236	152363	>>>	127	28	>>>	A1E_00670
3	A1E_02840	<<<	218	594155	594242	>>>	87	426	>>>	A1E_02845
4	A1E_05200	<<<	567	1064658	1064779	>>>	121	327	<<<	A1E_05205
5	A1E_05585	<<<	6	1147042	1147271	>>>	229	-1	<<<	A1E_05590
6	A1E_04445	<<<	173	893435	893619	<<<	184	205	>>>	A1E_04450
7	A1E_03540	>>>	259	733592	733663	<<<	71	376	>>>	A1E_03545
8	A1E_02545	<<<	210	549039	549095	<<<	56	180	>>>	A1E_02550
9	A1E_00080	>>>	150	20294	20590	<<<	296	1321	<<<	A1E_00085
10	A1E_00080	>>>	656	20800	20872	>>>	72	1039	<<<	A1E_00085
11	A1E_00080	>>>	630	20774	20903	<<<	129	1008	<<<	A1E_00085
12	A1E_00145	>>>	34	32661	32776	>>>	115	255	>>>	A1E_00150
13	A1E_00225	>>>	42	53352	53427	>>>	75	71	>>>	A1E_00230
14	A1E_00240	>>>	37	56312	56357	>>>	45	169	<<<	A1E_00245
15	A1E_00240	>>>	164	56439	56473	<<<	34	53	<<<	A1E_00245
16	A1E_00320	>>>	178	76723	76815	<<<	92	92	<<<	A1E_00325
17	A1E_00320	>>>	213	76758	76815	<<<	57	92	<<<	A1E_00325
19	A1E_t05688	>>>	88	109997	110293	>>>	296	27	>>>	A1E_00485
20	A1E_00560	>>>	86	133502	133636	>>>	134	35	>>>	A1E_00565
21	A1E_00665	>>>	1	152236	152390	>>>	154	1	>>>	A1E_00670
22	A1E_00710	>>>	13	159214	159500	>>>	286	402	>>>	A1E_00715
23	A1E_00810	>>>	0	185067	185276	<<<	209	398	<<<	A1E_t05692
24	A1E_01135	>>>	27	254109	254152	>>>	43	542	<<<	A1E_01140
25	A1E_01135	>>>	469	254551	254619	>>>	68	75	<<<	A1E_01140

26	A1E_01240	>>>	1	281996	282230	>>>	234	69	>>>	A1E_01245
27	A1E_01275	<<<	142	290911	291400	<<<	489	88	<<<	A1E_01280
28	A1E_01360	>>>	530	308971	309060	<<<	89	7	<<<	A1E_01365
29	A1E_01860	>>>	40	416032	416095	>>>	63	420	<<<	A1E_01865
30	A1E_01890	>>>	83	420872	420998	<<<	126	28	<<<	A1E_01895
31	A1E_t05700	<<<	865	429507	429628	<<<	121	1000	<<<	A1E_01940
32	A1E_02320	>>>	1	508611	508758	>>>	147	106	>>>	A1E_02325
33	A1E_02365	<<<	73	519574	519654	<<<	80	5	<<<	A1E_02370
34	A1E_02505	<<<	620	544273	544363	<<<	90	687	<<<	A1E_02510
35	A1E_03180	>>>	10	662831	662956	>>>	125	59	>>>	A1E_03185
36	A1E_04040	<<<	226	824941	825442	<<<	501	153	<<<	A1E_04045
37	A1E_04105	>>>	377	838802	838863	<<<	61	1149	<<<	A1E_04110
38	A1E_04420	<<<	1213	888649	888711	<<<	62	29	<<<	A1E_04425
39	A1E_04475	>>>	936	901046	901114	>>>	68	274	>>>	A1E_04480
40	A1E_04585	>>>	622	927464	927532	>>>	68	265	>>>	A1E_04590
41	A1E_04765	<<<	96	967562	967648	<<<	86	6	<<<	A1E_04770
42	A1E_04880	>>>	99	993954	994062	>>>	108	1663	>>>	A1E_04910
43	A1E_05125	<<<	128	1052138	1052680	<<<	542	1	<<<	A1E_05130
44	A1E_05305	>>>	-13	1091383	1091727	<<<	344	129	<<<	A1E_05310
45	A1E_05315	<<<	139	1093054	1093109	<<<	55	61	<<<	A1E_05320
46	A1E_05555	<<<	111	1141524	1141682	<<<	158	39	<<<	A1E_05560
47	A1E_02195	>>>	142	481263	481385	>>>	122	14	>>>	A1E_02200
48	A1E_04750	<<<	175	963690	963729	>>>	39	351	>>>	A1E_04755

Rickettsia conorii SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RC0097	>>>	116	92575	92721	>>>	146	437	<<<	RC0098
2	RC0150	>>>	1	151550	151679	>>>	129	564	>>>	RC0151
3	RC0375	<<<	179	375697	375836	>>>	139	130	<<<	RC0376
4	RC0448	<<<	1	442624	442681	>>>	57	49	<<<	RC0449
5	RC0593	<<<	16	580992	581052	>>>	60	138	>>>	RC0594
6	RC0633	>>>	1	619218	619310	>>>	92	22	>>>	RC0634
7	RC0782	<<<	246	739506	739595	>>>	89	580	>>>	RC0783
8	RC1056	<<<	1	980514	980727	>>>	213	1246	>>>	RC1057
9	RC1095	<<<	1	1020872	1020952	>>>	80	183	<<<	RC1096
10	RC1337	>>>	-24	1241587	1241949	<<<	362	131	<<<	RC1338
11	RC1120	>>>	653	1042270	1042553	<<<	283	9	>>>	RC1121
12	RC1102	<<<	-5	1026622	1026664	<<<	42	548	>>>	RC1103
13	RC0894	>>>	110	841793	841913	<<<	120	13	>>>	RC0895
14	RC0826	>>>	120	780179	780704	<<<	525	613	>>>	RCRNA22
15	RC0701	>>>	21	674690	675153	<<<	463	80	>>>	RC0702
16	RC0481	<<<	273	477661	477738	<<<	77	73	<<<	RC0482
17	RC0351	>>>	333	354359	354432	<<<	73	3	>>>	RC0352
18	RC0219	>>>	417	229291	229521	<<<	230	0	<<<	RC0220
19	RC0031	<<<	324	32400	32494	<<<	94	5	<<<	RC0032
20	RC0044	>>>	55	42041	42249	<<<	208	127	<<<	RC0045
21	RC0059	>>>	116	52581	52642	<<<	61	62	<<<	RC0060
22	RC0068	>>>	448	61996	62026	>>>	30	203	>>>	RC0069
23	RC0073	>>>	42	67028	67111	>>>	83	65	>>>	RC0074

24	RC0092	>>>	44	85536	85667	>>>	131	-10	<<<	RC0093
25	RC0097	>>>	338	92797	93067	<<<	270	91	<<<	RC0098
26	RC0109	>>>	76	104284	104811	<<<	527	297	<<<	RC0110
27	RC0112	>>>	69	113313	113378	>>>	65	61	>>>	RC0113
28	RC0141	>>>	1	134587	134709	>>>	122	-7	>>>	RC0142
29	RC0166	>>>	2	166510	166646	>>>	136	37	>>>	RC0167
30	RC0171	>>>	1	171221	171361	>>>	140	18	>>>	RC0172
31	RC0176	>>>	1	175608	175778	>>>	170	98	>>>	RC0177
32	RC0180	>>>	676	178909	178963	>>>	54	210	>>>	RC0181
33	RC0181	>>>	16	183312	183371	>>>	59	181	>>>	RC0182
34	RC0182	>>>	1	187673	188079	>>>	406	-1	<<<	RC0183
35	RC0188	<<<	245	196978	197113	<<<	135	0	<<<	RC0189
36	RC0219	>>>	592	229466	229521	<<<	55	0	<<<	RC0220
37	RC0226	<<<	78	233892	233977	<<<	85	0	<<<	RC0227
38	RC0228	<<<	40	238429	238549	<<<	120	15	<<<	RC0229
39	RC0241	>>>	24	251936	252254	<<<	318	0	<<<	RC0242
40	RC0241	>>>	1	251913	251998	>>>	85	256	<<<	RC0242
41	RC0241	>>>	91	252003	252121	>>>	118	133	<<<	RC0242
42	RC0241	>>>	151	252063	252254	<<<	191	0	<<<	RC0242
43	RC0243	>>>	108	253896	253963	<<<	67	122	<<<	RC0244
44	RC0243	>>>	115	253903	253954	>>>	51	131	<<<	RC0244
45	RC0257	>>>	400	262951	263040	>>>	89	85	<<<	RC0258
46	RC0282	<<<	129	287255	287339	<<<	84	0	<<<	RC0283
47	RC0332	>>>	13	334165	334271	>>>	106	123	>>>	RC0333
48	RC0332	>>>	13	334165	334313	>>>	148	81	>>>	RC0333
49	RC0335	>>>	25	337549	337665	<<<	116	0	<<<	RC0336

50	RC0342	>>>	19	344734	344857	>>>	123	107	>>>	RC0343
51	RC0352	>>>	19	355010	355142	>>>	132	24	>>>	RC0353
52	RC0352	>>>	19	355010	355164	>>>	154	2	>>>	RC0353
53	RC0359	>>>	45	361167	361287	>>>	120	324	>>>	RC0360
54	RC0388	>>>	1	388642	388729	>>>	87	95	>>>	RC0389
55	RC0412	>>>	45	414336	414527	>>>	191	195	>>>	RC0413
56	RC0414	>>>	0	415202	415638	<<<	436	0	<<<	RC0415
57	RC0414	>>>	32	415234	415286	>>>	52	352	<<<	RC0415
58	RC0431	<<<	81	429228	429315	<<<	87	18	<<<	RC0432
59	RC0447	>>>	11	441137	441225	>>>	88	53	<<<	RC0448
60	RC0447	>>>	41	441167	441278	<<<	111	0	<<<	RC0448
61	RC0455	>>>	6	450533	450705	>>>	172	421	>>>	RC0456
62	RC0467	<<<	1199	464303	464511	<<<	208	0	<<<	RC0468
63	RC0493	>>>	6	487697	487818	>>>	121	69	<<<	RC0494
64	RC0493	>>>	6	487697	487800	>>>	103	87	<<<	RC0494
65	RC0497	>>>	84	491842	491968	>>>	126	594	>>>	RC0498
66	RC0506	>>>	1	500186	500284	>>>	98	38	>>>	RC0507
67	RC0522	<<<	94	512453	512522	>>>	69	47	<<<	RC0523
68	RC0546	>>>	1	536256	536393	>>>	137	77	>>>	RC0547
69	RC0557	>>>	44	550166	550215	>>>	49	84	>>>	RC0558
70	RC0557	>>>	44	550166	550259	>>>	93	40	>>>	RC0558
71	RC0563	>>>	654	557086	557184	>>>	98	104	<<<	RC0564
72	RC0563	>>>	654	557086	557164	>>>	78	124	<<<	RC0564
73	RC0563	>>>	688	557120	557217	<<<	97	71	<<<	RC0564
74	RC0583	>>>	199	572072	572162	>>>	90	64	>>>	RC0584
75	RC0596	>>>	1	583592	583719	>>>	127	1	>>>	RC0597

76	RC0602	<<<	67	588314	588400	<<<	86	0	<<<	RC0603
77	RC0605	<<<	31	592014	592174	<<<	160	400	<<<	RC0606
78	RC0611	<<<	91	596403	596486	<<<	83	0	<<<	RC0612
79	RC0617	>>>	15	602320	602469	>>>	149	33	>>>	RC0618
80	RC0621	>>>	23	606064	606109	>>>	45	84	>>>	RC0622
81	RC0627	>>>	1	611703	611840	>>>	137	37	<<<	RC0628
82	RC0629	>>>	181	615513	615644	>>>	131	370	>>>	RC0630
83	RC0629	>>>	181	615513	615614	>>>	101	400	>>>	RC0630
84	RC0633	>>>	1	619218	619333	>>>	115	-1	>>>	RC0634
85	RC0645	>>>	193	626412	626527	>>>	115	204	>>>	RC0646
86	RC0667	>>>	1	651312	651408	>>>	96	20	<<<	RC0668
87	RC0679	<<<	91	659379	659492	<<<	113	41	<<<	RCRNA15
88	RCRNA15	<<<	115	659734	659866	<<<	132	0	<<<	RC0680
89	RC0693	>>>	6	669401	669607	>>>	206	10	>>>	RC0694
90	RC0713	>>>	63	685895	685978	>>>	83	82	<<<	RC0714
91	RC0811	>>>	1	766677	766745	>>>	68	51	>>>	RC0812
92	RC0822	<<<	54	777815	777918	<<<	103	0	<<<	RC0823
93	RC0826	>>>	20	780079	780235	>>>	156	1082	>>>	RCRNA22
94	RC0850	>>>	111	801483	801594	<<<	111	0	<<<	RC0851
95	RC0851	<<<	277	802832	802912	<<<	80	316	<<<	RC0852
96	RC0911	<<<	65	857009	857131	<<<	122	19	<<<	RC0912
97	RC0937	>>>	250	877391	877540	>>>	149	11	<<<	RC0938
98	RC0956	>>>	450	896283	896383	>>>	100	148	>>>	RC0957
99	RC0956	>>>	450	896283	896344	>>>	61	187	>>>	RC0957
100	RC0956	>>>	450	896283	896501	>>>	218	30	>>>	RC0957
101	RC0965	>>>	3	906730	906975	>>>	245	49	>>>	RC0966

102	RC0967	>>>	40	909140	909208	>>>	68	849	<<<	RC0968
103	RC0967	>>>	353	909453	909709	<<<	256	348	<<<	RC0968
104	RC0974	>>>	92	914960	915021	<<<	61	0	<<<	RC0975
105	RC0979	>>>	729	919114	919257	>>>	143	1006	<<<	RCRNA27
106	RC0979	>>>	1647	920032	920121	>>>	89	142	<<<	RCRNA27
107	RC0979	>>>	1635	920020	920240	<<<	220	23	<<<	RCRNA27
108	RC1041	>>>	27	966835	966988	>>>	153	77	<<<	RC1042
109	RC1041	>>>	140	966948	967049	<<<	101	16	<<<	RC1042
110	RC1048	>>>	-4	974182	974336	<<<	154	0	<<<	RC1049
111	RC1073	>>>	7	994661	994792	>>>	131	163	>>>	RC1074
112	RC1074	>>>	586	995791	995852	>>>	61	18	>>>	RC1075
113	RC1084	<<<	499	1007117	1007168	<<<	51	19	<<<	RC1085
114	RC1090	>>>	1	1016864	1017022	>>>	158	-41	<<<	RC1091
115	RCRNA33	>>>	4	1040304	1040654	>>>	350	22	<<<	RC1119
116	RC1120	>>>	292	1041909	1042044	>>>	135	518	>>>	RC1121
117	RC1120	>>>	292	1041909	1042014	>>>	105	548	>>>	RC1121
118	RC1123	>>>	527	1043770	1044310	<<<	540	0	<<<	RC1124
119	RC1138	>>>	57	1052324	1052449	<<<	125	339	<<<	RC1139
120	RC1186	>>>	123	1087009	1087096	>>>	87	70	>>>	RC1187
121	RC1228	>>>	2	1133330	1133619	>>>	289	186	<<<	RC1229
122	RC1254	>>>	1	1159676	1159876	>>>	200	-57	<<<	RC1255
123	RC1254	>>>	1	1159676	1159833	>>>	157	-14	<<<	RC1255
124	RC1280	>>>	327	1188699	1188764	<<<	65	0	<<<	RC1281
125	RC1280	>>>	6	1188378	1188751	>>>	373	13	<<<	RC1281
126	RC1303	>>>	75	1207263	1207375	>>>	112	747	>>>	RC1304
127	RC1323	<<<	136	1224816	1224892	<<<	76	1	<<<	RC1324

128	RCRNA38	<<<	91	1256610	1256702	<<<	92	0	<<<	RC1359
129	RC1365	>>>	103	1261170	1261383	>>>	213	282	<<<	RC1366
130	RC1369	>>>	1	1264399	1264469	>>>	70	183	<<<	RC1370
131	RC1369	>>>	88	1264486	1264558	>>>	72	94	<<<	RC1370
132	RC1084	<<<	244	1006862	1007152	>>>	290	35	<<<	RC1085
133	RC0128	>>>	47	124698	124792	<<<	94	14	>>>	RC0129
134	RC0141	>>>	1	134587	134679	>>>	92	23	>>>	RC0142
135	RC0214	>>>	212	225148	225199	>>>	51	436	<<<	RC0215
136	RC0234	>>>	221	246417	246557	<<<	140	452	>>>	RC0235
137	RC0330	>>>	116	331482	331585	<<<	103	0	>>>	RC0331
138	RC0359	>>>	45	361167	361250	>>>	83	361	>>>	RC0360
139	RC0472	<<<	113	470814	470946	<<<	132	39	<<<	RC0473
140	RC0493	>>>	30	487721	487873	<<<	152	14	<<<	RC0494
141	RC0527	>>>	87	518306	518349	<<<	43	33	>>>	RC0528
142	RC0779	<<<	175	737032	737106	<<<	74	2	<<<	RC0780
143	RC0818	<<<	16	775091	775197	<<<	106	92	<<<	RC0819
144	RC0847	<<<	1	795966	796035	>>>	69	106	<<<	RC0848
145	RC0102	>>>	137	97732	97769	>>>	37	133	<<<	RC0103

Rickettsia felis SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RF_0052	<<<	79	53567	53714	>>>	147	832	<<<	RF_0053
2	RF_0185	<<<	162	203281	203376	>>>	95	80	>>>	RF_0186
3	RF_0378	>>>	922	402257	402324	>>>	67	589	>>>	RF_0379
4	RF_0693	>>>	41	751761	751922	>>>	161	146	>>>	RF_0694
5	RF_1067	>>>	9	1129055	1129552	>>>	497	-56	<<<	RF_1068
6	RF_1183	<<<	1	1252953	1253070	>>>	117	64	>>>	RF_1184
7	RF_1334	>>>	211	1420894	1421010	<<<	116	30	>>>	RF_1335
8	RF_0929	>>>	353	989853	989987	<<<	134	169	>>>	RF_0930
9	RF_0780	>>>	116	836630	836706	<<<	76	146	>>>	RF_0781
10	RF_0735	<<<	576	799527	799618	<<<	91	461	>>>	RF_0736
11	RF_0684	>>>	247	731985	732070	<<<	85	67	>>>	RF_0685
12	RF_0470	>>>	261	501131	501267	<<<	136	0	>>>	RF_0471
13	RF_0244	>>>	236	272732	272875	<<<	143	0	>>>	RF_0245
14	RF_0009	>>>	36	10477	10617	>>>	140	328	>>>	RF_0010
15	RF_0009	>>>	36	10477	10585	>>>	108	360	>>>	RF_0010
16	RF_0010	>>>	155	11404	11513	>>>	109	87	<<<	RF_0011
17	RF_0010	>>>	224	11473	11567	<<<	94	33	<<<	RF_0011
18	RF_0034	<<<	103	36802	36894	<<<	92	8	<<<	RF_0035
19	RF_0052	<<<	292	53780	54084	<<<	304	462	<<<	RF_0053
20	RF_0052	<<<	914	54402	54459	<<<	57	87	<<<	RF_0053
21	RF_0057	>>>	39	59191	59302	>>>	111	154	<<<	RF_0058
22	RF_0070	>>>	5	77036	77128	>>>	92	104	>>>	RF_0071
23	RF_RNA04	>>>	48	89616	89895	>>>	279	84	>>>	RF_0085

24	RF_0091	>>>	2	102909	103010	>>>	101	124	>>>	RF_0092
25	RF_0103	<<<	174	116606	116915	<<<	309	14	<<<	RF_0104
26	RF_0111	>>>	72	125658	125758	<<<	100	92	<<<	RF_0112
27	RF_0111	>>>	184	125770	125826	<<<	56	24	<<<	RF_0112
28	RF_0140	<<<	201	158827	158903	<<<	76	0	<<<	RF_0141
29	RF_0159	>>>	23	175827	175862	>>>	35	149	>>>	RF_0160
30	RF_0182	>>>	1	201331	201411	>>>	80	79	>>>	RF_0183
31	RF_0184	>>>	1	202238	202303	>>>	65	284	<<<	RF_0185
32	RF_0204	>>>	88	219599	219643	>>>	44	332	>>>	RF_0205
33	RF_0204	>>>	88	219599	219675	>>>	76	300	>>>	RF_0205
34	RF_0205	>>>	1	224942	225032	>>>	90	564	>>>	RF_0206
35	RF_0205	>>>	387	225328	225415	>>>	87	181	>>>	RF_0206
37	RF_0262	>>>	5	291474	291580	>>>	106	538	>>>	RF_0263
38	RF_0265	>>>	17	294743	294804	>>>	61	86	<<<	RF_0266
39	RF_0265	>>>	17	294743	294847	>>>	104	43	<<<	RF_0266
40	RF_0265	>>>	17	294743	294829	>>>	86	61	<<<	RF_0266
41	RF_0265	>>>	59	294785	294886	<<<	101	4	<<<	RF_0266
42	RF_RNA11	>>>	185	319853	319947	>>>	94	923	>>>	RF_0305
43	RF_RNA11	>>>	366	320034	320088	>>>	54	782	>>>	RF_0305
44	RF_0315	>>>	200	331166	331399	>>>	233	151	>>>	RF_0316
45	RF_0316	>>>	19	331828	332206	>>>	378	852	>>>	RF_0317
46	RF_RNA12	>>>	22	344467	344607	>>>	140	98	>>>	RF_RNA13
47	RF_0332	>>>	35	352842	352894	>>>	52	121	>>>	RF_0333
48	RF_0332	>>>	35	352842	352995	>>>	153	20	>>>	RF_0333
49	RF_0346	>>>	126	367530	367657	<<<	127	241	<<<	RF_0347
50	RF_0346	>>>	7	367411	367570	>>>	159	328	<<<	RF_0347

51	RF_0346	>>>	273	367677	367750	>>>	73	148	<<<	RF_0347
52	RF_0351	>>>	20	370042	370293	>>>	251	140	>>>	RF_0352
53	RF_0357	>>>	87	374577	374728	>>>	151	142	>>>	RF_0358
54	RF_0378	>>>	21	401356	401467	>>>	111	1446	>>>	RF_0379
55	RF_0409	>>>	183	430050	430157	>>>	107	1103	>>>	RF_0410
56	RF_0411	>>>	62	432986	433077	<<<	91	13	<<<	RF_0412
57	RF_0413	<<<	279	435295	435352	<<<	57	2	<<<	RF_0414
58	RF_0421	<<<	33	446138	446311	<<<	173	340	<<<	RF_0422
59	RF_0421	<<<	152	446257	446311	<<<	54	340	<<<	RF_0422
60	RF_0421	<<<	425	446530	446638	<<<	108	13	<<<	RF_0422
61	RF_0426	>>>	26	452975	453032	>>>	57	208	>>>	RF_0427
62	RF_0432	>>>	143	460511	460551	>>>	40	388	>>>	RF_0433
63	RF_0432	>>>	116	460484	460527	>>>	43	412	>>>	RF_0433
64	RF_0441	<<<	104	469696	469850	>>>	154	425	<<<	RF_0442
65	RF_0470	>>>	28	500898	501024	>>>	126	243	>>>	RF_0471
66	RF_0470	>>>	28	500898	501003	>>>	105	264	>>>	RF_0471
67	RF_0470	>>>	132	501002	501077	>>>	75	190	>>>	RF_0471
68	RF_0470	>>>	132	501002	501047	>>>	45	220	>>>	RF_0471
69	RF_0492	>>>	140	525646	525757	<<<	111	146	<<<	RF_0493
70	RF_0492	>>>	285	525791	525955	>>>	164	-52	<<<	RF_0493
71	RF_0506	>>>	28	540947	541097	>>>	150	350	<<<	RF_0507
72	RF_0506	>>>	28	540947	541067	>>>	120	380	<<<	RF_0507
73	RF_0510	<<<	78	546179	546283	<<<	104	18	<<<	RF_0511
74	RF_0535	>>>	6	572179	572354	>>>	175	96	>>>	RF_0536
75	RF_0544	<<<	200	581686	581888	<<<	202	32	<<<	RF_0545
76	RF_0548	<<<	-5	585491	585762	<<<	271	941	<<<	RF_0549

77	RF_0548	<<<	840	586336	586573	<<<	237	130	<<<	RF_0549
78	RF_0552	>>>	425	590379	590412	<<<	33	452	<<<	RF_0553
79	RF_0574	>>>	224	611046	611284	<<<	238	3	<<<	RF_0575
80	RF_0574	>>>	247	611069	611181	>>>	112	106	<<<	RF_0575
81	RF_0577	>>>	691	614952	615192	>>>	240	311	>>>	RF_0578
82	RF_0577	>>>	691	614952	615427	>>>	475	76	>>>	RF_0578
83	RF_0591	>>>	85	635163	635286	>>>	123	81	<<<	RF_0592
84	RF_0591	>>>	168	635246	635364	<<<	118	3	<<<	RF_0592
85	RF_0626	<<<	42	670494	670633	<<<	139	2	<<<	RF_0627
86	RF_0638	>>>	4	683524	683623	>>>	99	34	<<<	RF_0639
87	RF_0638	>>>	4	683524	683602	>>>	78	55	<<<	RF_0639
88	RF_0638	>>>	38	683558	683656	<<<	98	1	<<<	RF_0639
89	RF_0646	>>>	28	693172	693381	>>>	209	-9	>>>	RF_0647
90	RF_0674	<<<	18	720914	720948	>>>	34	268	<<<	RF_0675
91	RF_0674	<<<	46	720942	721030	>>>	88	186	<<<	RF_0675
92	RF_0674	<<<	64	720960	721044	<<<	84	172	<<<	RF_0675
93	RF_0674	<<<	221	721117	721190	<<<	73	26	<<<	RF_0675
94	RF_0674	<<<	46	720942	721161	>>>	219	55	<<<	RF_0675
95	RF_0696	>>>	1	754992	755129	>>>	137	-21	>>>	RF_0697
96	RF_0713	>>>	17	772969	773056	>>>	87	81	<<<	RF_0714
97	RF_0725	>>>	5	788304	788425	>>>	121	86	<<<	RF_0726
98	RF_0725	>>>	66	788365	788494	<<<	129	17	<<<	RF_0726
99	RF_0730	>>>	1	792699	792946	>>>	247	-25	>>>	RF_0731
100	RF_0764	>>>	106	821918	821979	>>>	61	418	>>>	RF_0765
101	RF_0806	>>>	16	855022	855166	>>>	144	100	>>>	RF_0807
102	RF_0808	>>>	322	857628	857779	>>>	151	-22	<<<	RF_0809

103	RF_0813	>>>	27	862241	862355	>>>	114	201	>>>	RF_0814
104	RF_0826	>>>	183	873945	874046	<<<	101	2	<<<	RF_0827
105	RF_0828	>>>	59	875243	875341	<<<	98	162	<<<	RF_0829
106	RF_0828	>>>	14	875198	875283	>>>	85	220	<<<	RF_0829
107	RF_0841	<<<	159	891113	891324	<<<	211	109	<<<	RF_0842
108	RF_0847	<<<	265	895895	896010	<<<	115	9	<<<	RF_0848
109	RF_0855	<<<	54	907702	907834	<<<	132	0	<<<	RF_0856
110	RF_0859	>>>	1	912621	912873	>>>	252	-9	>>>	RF_0860
111	RF_0870	<<<	244	925942	926037	<<<	95	180	<<<	RF_0871
112	RF_0878	>>>	72	935933	936062	>>>	129	609	>>>	RF_0879
113	RF_0879	>>>	15	937029	937099	<<<	70	281	<<<	RF_0880
114	RF_0887	>>>	127	950375	950500	>>>	125	439	<<<	RF_0888
115	RF_0887	>>>	189	950437	950634	<<<	197	305	<<<	RF_0888
116	RF_0887	>>>	127	950375	950479	>>>	104	460	<<<	RF_0888
117	RF_0887	>>>	230	950478	950634	<<<	156	305	<<<	RF_0888
118	RF_0887	>>>	127	950375	950538	>>>	163	401	<<<	RF_0888
119	RF_0887	>>>	127	950375	950662	>>>	287	277	<<<	RF_0888
120	RF_0887	>>>	340	950588	950634	<<<	46	305	<<<	RF_0888
121	RF_0887	>>>	127	950375	950632	>>>	257	307	<<<	RF_0888
122	RF_0929	>>>	164	989664	989810	>>>	146	346	>>>	RF_0930
123	RF_0929	>>>	402	989902	989976	>>>	74	180	>>>	RF_0930
124	RF_0939	<<<	103	996968	997098	<<<	130	5	<<<	RF_0940
125	RF_0940	<<<	142	998116	998150	<<<	34	168	<<<	RF_0941
126	RF_0979	>>>	151	1038529	1038694	<<<	165	3	<<<	RF_0980
127	RF_0990	>>>	17	1053533	1053859	>>>	326	131	>>>	RF_0991
128	RF_0995	>>>	64	1057788	1057929	>>>	141	79	>>>	RF_RNA30

129	RF_0995	>>>	64	1057788	1057897	>>>	109	111	>>>	RF_RNA30
130	RF_1018	>>>	249	1079597	1079706	<<<	109	4	<<<	RF_1019
131	RF_1036	<<<	65	1099649	1099903	<<<	254	22	<<<	RF_1037
132	RF_1045	<<<	186	1106205	1106307	<<<	102	27	<<<	RF_1046
133	RF_1074	>>>	38	1142079	1142314	>>>	235	3	<<<	RF_1075
134	RF_1093	>>>	23	1163897	1164049	>>>	152	-3	>>>	RF_1094
135	RF_1098	>>>	321	1168483	1168527	<<<	44	367	<<<	RF_1099
136	RF_1098	>>>	295	1168457	1168543	>>>	86	351	<<<	RF_1099
137	RF_1098	>>>	371	1168533	1168646	<<<	113	248	<<<	RF_1099
138	RF_1098	>>>	359	1168521	1168575	>>>	54	319	<<<	RF_1099
139	RF_1098	>>>	464	1168626	1168707	>>>	81	187	<<<	RF_1099
140	RF_1098	>>>	464	1168626	1168686	>>>	60	208	<<<	RF_1099
141	RF_1098	>>>	484	1168646	1168714	<<<	68	180	<<<	RF_1099
142	RF_1105	>>>	83	1174383	1174427	>>>	44	322	<<<	RF_1106
143	RF_1105	>>>	87	1174387	1174597	<<<	210	152	<<<	RF_1106
144	RF_1121	>>>	21	1190664	1190784	>>>	120	84	>>>	RF_RNA32
145	RF_1126	>>>	23	1195161	1195250	>>>	89	203	>>>	RF_1127
146	RF_1126	>>>	5	1195143	1195412	>>>	269	41	>>>	RF_1127
147	RF_1134	>>>	1	1202319	1202393	>>>	74	396	>>>	RF_1135
148	RF_1144	>>>	21	1213996	1214097	<<<	101	220	<<<	RF_1145
149	RF_1144	>>>	17	1213992	1214062	>>>	70	255	<<<	RF_1145
150	RF_1144	>>>	166	1214141	1214317	<<<	176	0	<<<	RF_1145
151	RF_1146	<<<	250	1223011	1223361	<<<	350	342	<<<	RF_1147
152	RF_1146	<<<	670	1223431	1223692	<<<	261	11	<<<	RF_1147
153	RF_1186	<<<	176	1255047	1255177	<<<	130	11	<<<	RF_1187
154	RF_1230	<<<	87	1295321	1295652	<<<	331	95	<<<	RF_1231

155	RF_RNA34	<<<	231	1299279	1299415	<<<	136	105	<<<	RF_1234
156	RF_RNA34	<<<	410	1299458	1299508	<<<	50	12	<<<	RF_1234
157	RF_1239	>>>	2	1306083	1306296	>>>	213	87	>>>	RF_1240
158	RF_1243	>>>	140	1310473	1310592	<<<	119	9	<<<	RF_1244
159	RF_1249	>>>	37	1321868	1322031	>>>	163	43	>>>	RF_1250
160	RF_1260	<<<	115	1338282	1338344	<<<	62	7	<<<	RF_1261
161	RF_1285	>>>	40	1365709	1365971	>>>	262	28	>>>	RF_1286
162	RF_1293	>>>	56	1378419	1378503	<<<	84	17	<<<	RF_1294
163	RF_1300	<<<	129	1386973	1387022	<<<	49	231	<<<	RF_1301
165	RF_1312	>>>	140	1400537	1400817	>>>	280	9	<<<	RF_1313
166	RF_1312	>>>	366	1400763	1400811	<<<	48	15	<<<	RF_1313
167	RF_1322	>>>	26	1407546	1407669	>>>	123	85	>>>	RF_1323
168	RF_1334	>>>	41	1420724	1420872	>>>	148	168	>>>	RF_1335
169	RF_1334	>>>	41	1420724	1420851	>>>	127	189	>>>	RF_1335
170	RF_1334	>>>	41	1420724	1420949	>>>	225	91	>>>	RF_1335
171	RF_1335	>>>	1	1421462	1421547	>>>	85	81	>>>	RF_1336
172	RF_1336	>>>	7	1422974	1423101	>>>	127	86	<<<	RF_1337
173	RF_1336	>>>	94	1423061	1423150	<<<	89	37	<<<	RF_1337
174	RF_1338	<<<	245	1425945	1426055	<<<	110	1	<<<	RF_1339
175	RF_1353	>>>	101	1443167	1443217	<<<	50	730	<<<	RF_1354
176	RF_1353	>>>	231	1443297	1443367	<<<	70	580	<<<	RF_1354
177	RF_1353	>>>	213	1443279	1443337	>>>	58	610	<<<	RF_1354
178	RF_1397	<<<	469	1482020	1482147	<<<	127	112	<<<	RF_1398
179	RF_0986	>>>	14	1046043	1046148	>>>	105	24	>>>	RF_0987
180	RF_RNA30	>>>	121	1058207	1058346	<<<	139	134	>>>	RF_0996
181	RF_0009	>>>	161	10602	10803	>>>	201	142	>>>	RF_0010

182	RF_0205	>>>	1	224942	225011	>>>	69	585	>>>	RF_0206
183	RF_0205	>>>	173	225114	225154	<<<	40	442	>>>	RF_0206
184	RF_0212	<<<	200	236710	236755	<<<	45	338	<<<	RF_0213
185	RF_0238	>>>	110	267037	267174	<<<	137	47	<<<	RF_0239
186	RF_0523	>>>	167	558282	558325	<<<	43	47	>>>	RF_0524
187	RF_0552	>>>	752	590706	590851	<<<	145	13	<<<	RF_0553
188	RF_0627	<<<	18	672253	672325	>>>	72	158	<<<	RF_0628
189	RF_0723	<<<	53	783050	783249	>>>	199	238	>>>	RF_0724
190	RF_0873	<<<	16	931019	931125	<<<	106	3	<<<	RF_0874

Rickettsia heilongjiangensis SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	Rh054_02570	<<<	1	449233	449290	>>>	57	49	<<<	Rh054_02575
2	Rh054_03340	<<<	16	587636	587696	>>>	60	136	>>>	Rh054_03345
3	Rh054_03405	<<<	1	597715	597979	>>>	264	61	<<<	Rh054_03406
4	Rh054_03565	>>>	1	626718	626810	>>>	92	22	>>>	Rh054_03570
5	Rh054_05285	<<<	1	916959	917031	>>>	72	59	>>>	Rh054_05290
6	Rh054_06055	<<<	443	1037148	1037228	>>>	80	183	<<<	Rh054_06060
7	Rh054_07320	>>>	2	1257718	1257823	<<<	105	26	>>>	Rh054_07325
8	Rh054_06910	<<<	501	1182979	1183047	<<<	68	138	<<<	Rh054_06915
9	Rh054_04935	>>>	110	852825	852945	<<<	120	13	>>>	Rh054_04940
10	Rh054_04555	<<<	2245	788214	788294	<<<	80	13	<<<	Rh054_04570
11	Rh054_04125	>>>	687	716662	716979	<<<	317	0	<<<	Rh054_04140
12	Rh054_03870	>>>	161	679380	679468	<<<	88	337	>>>	Rh054_03885
13	Rh054_03800	<<<	594	670737	670828	<<<	91	244	>>>	Rh054_03805
14	Rh054_02360	>>>	452	420958	421318	<<<	360	0	>>>	Rh054_02374
15	Rh054_02115	<<<	1064	377206	377413	<<<	207	325	>>>	Rh054_02130
16	Rh054_01990	>>>	143	360642	360716	<<<	74	1	>>>	Rh054_01995
17	Rh054_01490	>>>	262	268483	268653	<<<	170	62	<<<	Rh054_01495
18	Rh054_01300	>>>	-25	235383	235530	<<<	147	281	<<<	Rh054_01305
19	Rh054_00825	>>>	84	137268	137373	<<<	105	25	>>>	Rh054_00830
20	Rh054_00115	>>>	1	25942	26057	>>>	115	-4	>>>	Rh054_t07484
21	Rh054_00180	<<<	18	31816	31907	<<<	91	7	<<<	Rh054_00185
22	Rh054_00360	>>>	116	56909	56970	<<<	61	62	<<<	Rh054_00365
23	Rh054_00450	>>>	42	72004	72087	>>>	83	64	>>>	Rh054_00455
24	Rh054_00575	>>>	4	90784	90915	>>>	131	-10	<<<	Rh054_00580

25	Rh054_00600	>>>	9	95325	95465	>>>	140	40	>>>	Rh054_00605
26	Rh054_00605	>>>	89	97809	97853	>>>	44	320	<<<	Rh054_00610
27	Rh054_00630	>>>	16	102624	102772	>>>	148	153	<<<	Rh054_00635
28	Rh054_00685	>>>	56	118561	118617	>>>	56	70	>>>	Rh054_00690
29	Rh054_00830	>>>	1	139818	139940	>>>	122	-7	>>>	Rh054_00835
30	Rh054_00985	>>>	8	171833	171967	>>>	134	33	>>>	Rh054_00990
31	Rh054_01015	>>>	1	176533	176673	>>>	140	18	>>>	Rh054_01020
32	Rh054_01040	>>>	1	180918	181088	>>>	170	98	>>>	Rh054_01045
33	Rh054_01065	>>>	16	188580	188639	>>>	59	183	>>>	Rh054_01070
34	Rh054_01070	>>>	1	192943	193341	>>>	398	-1	<<<	Rh054_01075
35	Rh054_01160	>>>	-4	212938	213082	<<<	144	231	<<<	Rh054_01165
36	Rh054_01180	<<<	85	217236	217287	<<<	51	102	<<<	Rh054_01185
37	Rh054_01330	<<<	57	239929	240035	<<<	106	0	<<<	Rh054_01335
38	Rh054_01415	>>>	1	257320	257405	>>>	85	347	<<<	Rh054_01420
39	Rh054_01415	>>>	24	257343	257752	<<<	409	0	<<<	Rh054_01420
40	Rh054_01415	>>>	241	257560	257752	<<<	192	0	<<<	Rh054_01420
41	Rh054_01415	>>>	1	257320	257618	>>>	298	134	<<<	Rh054_01420
42	Rh054_01425	>>>	115	259401	259452	>>>	51	93	<<<	Rh054_01430
43	Rh054_01425	>>>	108	259394	259460	<<<	66	85	<<<	Rh054_01430
44	Rh054_01490	>>>	328	268549	268634	>>>	85	81	<<<	Rh054_01495
45	Rh054_01900	>>>	14	340465	340556	>>>	91	31	>>>	Rh054_01905
46	Rh054_01915	>>>	25	343742	343860	<<<	118	0	<<<	Rh054_01920
47	Rh054_01945	>>>	10	350927	351056	>>>	129	107	>>>	Rh054_01950
48	Rh054_02245	>>>	1	396888	396999	>>>	111	39	>>>	Rh054_02250
49	Rh054_02355	>>>	102	419927	420062	>>>	135	164	>>>	Rh054_02360
50	Rh054_02455	<<<	81	435571	435659	<<<	88	18	<<<	Rh054_02460

51	Rh054_02565	>>>	74	447776	447882	<<<	106	5	<<<	Rh054_02570
52	Rh054_02565	>>>	33	447735	447834	>>>	99	53	<<<	Rh054_02570
53	Rh054_02595	>>>	1	452760	452953	>>>	193	119	<<<	Rh054_02600
54	Rh054_02620	>>>	6	457137	457308	>>>	171	85	>>>	Rh054_02625
55	Rh054_02695	<<<	78	471046	471255	<<<	209	0	<<<	Rh054_02700
56	Rh054_02755	>>>	1	482294	482439	>>>	145	107	<<<	Rh054_02760
57	Rh054_02850	>>>	84	498208	498333	>>>	125	493	>>>	Rh054_02855
58	Rh054_02900	>>>	1	506498	506596	>>>	98	39	>>>	Rh054_02905
59	Rh054_02925	>>>	21	512178	512323	>>>	145	51	<<<	Rh054_02930
60	Rh054_02925	>>>	104	512261	512366	<<<	105	8	<<<	Rh054_02930
61	Rh054_02970	>>>	265	517473	517789	<<<	316	22	<<<	Rh054_02975
62	Rh054_02970	>>>	274	517482	517675	>>>	193	136	<<<	Rh054_02975
63	Rh054_02970	>>>	274	517482	517645	>>>	163	166	<<<	Rh054_02975
64	Rh054_03010	>>>	9	526718	526945	>>>	227	-25	>>>	Rh054_03015
65	Rh054_03165	>>>	41	557249	557298	>>>	49	99	>>>	Rh054_03170
66	Rh054_03165	>>>	41	557249	557342	>>>	93	55	>>>	Rh054_03170
67	Rh054_03205	>>>	1	564231	564330	>>>	99	36	<<<	Rh054_03210
68	Rh054_03295	>>>	1	578572	578672	>>>	100	58	>>>	Rh054_03300
69	Rh054_03305	>>>	102	581973	582083	>>>	110	311	<<<	Rh054_03310
70	Rh054_03310	<<<	196	582750	582797	<<<	47	158	<<<	Rh054_03315
71	Rh054_03395	<<<	77	594867	594953	<<<	86	0	<<<	Rh054_03400
72	Rh054_03406	<<<	213	598584	598702	<<<	118	0	<<<	Rh054_03407
73	Rh054_03435	>>>	478	601861	601996	>>>	135	318	<<<	Rh054_03436
74	Rh054_03436	<<<	49	603402	603436	>>>	34	95	<<<	Rh054_03450
75	Rh054_03436	<<<	95	603448	603531	<<<	83	0	<<<	Rh054_03450
76	Rh054_03436	<<<	49	603402	603518	>>>	116	13	<<<	Rh054_03450

77	Rh054_03475	>>>	34	609850	609983	>>>	133	33	>>>	Rh054_03480
78	Rh054_03495	>>>	18	613575	613625	>>>	50	461	>>>	Rh054_03510
79	Rh054_03530	>>>	1	619358	619495	>>>	137	37	<<<	Rh054_03535
80	Rh054_03540	>>>	181	623168	623299	>>>	131	348	<<<	Rh054_03545
81	Rh054_03540	>>>	181	623168	623269	>>>	101	378	<<<	Rh054_03545
82	Rh054_03565	>>>	1	626718	626833	>>>	115	-1	>>>	Rh054_03570
83	Rh054_03596	>>>	427	633633	633748	>>>	115	8	>>>	Rh054_03620
84	Rh054_03700	<<<	37	647371	647495	<<<	124	0	<<<	Rh054_03705
85	Rh054_03795	>>>	75	667343	667457	<<<	114	48	<<<	Rh054_03800
86	Rh054_03815	>>>	4	673225	673306	>>>	81	166	>>>	Rh054_03820
87	Rh054_03910	>>>	2	683680	683800	>>>	120	22	>>>	Rh054_03915
88	Rh054_03985	>>>	17	695270	695350	>>>	80	170	>>>	Rh054_03990
89	Rh054_t07510	<<<	2	701587	701759	>>>	172	73	<<<	Rh054_04025
90	Rh054_t07510	<<<	116	701701	701832	<<<	131	0	<<<	Rh054_04025
91	Rh054_04095	>>>	80	711762	711894	>>>	132	79	>>>	Rh054_04100
92	Rh054_04105	>>>	69	712707	712801	>>>	94	34	<<<	Rh054_04110
93	Rh054_04105	>>>	69	712707	712857	>>>	150	-22	<<<	Rh054_04110
94	Rh054_04205	>>>	4	727411	727556	>>>	145	84	<<<	Rh054_04210
95	Rh054_04405	>>>	23	759805	759912	>>>	107	64	>>>	Rh054_04410
96	Rh054_04410	>>>	1	761463	761582	>>>	119	17	<<<	Rh054_04415
97	Rh054_04640	>>>	111	805776	805882	<<<	106	3	<<<	Rh054_04645
98	Rh054_04645	<<<	277	807123	807203	<<<	80	316	<<<	Rh054_04650
99	Rh054_04700	>>>	45	817682	817812	<<<	130	26	<<<	Rh054_04705
100	Rh054_04700	>>>	16	817653	817728	>>>	75	110	<<<	Rh054_04705
101	Rh054_04730	>>>	1	821939	822479	>>>	540	-10	<<<	Rh054_04735
102	Rh054_04915	<<<	241	849362	849575	<<<	213	352	>>>	Rh054_t07524

103	Rh054_04950	>>>	18	854741	854894	>>>	153	31	>>>	Rh054_04955
104	Rh054_05055	>>>	281	876271	876568	>>>	297	46	<<<	Rh054_05060
105	Rh054_05090	<<<	38	881848	881985	<<<	137	0	<<<	Rh054_05095
106	Rh054_05150	>>>	250	888475	888635	>>>	160	3	<<<	Rh054_05155
107	Rh054_05160	>>>	5	890467	890643	<<<	176	10	<<<	Rh054_05165
108	Rh054_05160	>>>	1	890463	890531	>>>	68	122	<<<	Rh054_05165
109	Rh054_05240	>>>	1	906049	906344	>>>	295	33	<<<	Rh054_05245
110	Rh054_05275	>>>	12	915748	915815	>>>	67	117	<<<	Rh054_05280
111	Rh054_05275	>>>	12	915748	915834	>>>	86	98	<<<	Rh054_05280
112	Rh054_05290	>>>	12	917973	918080	>>>	107	178	>>>	Rh054_05295
113	Rh054_05290	>>>	12	917973	918233	>>>	260	25	>>>	Rh054_05295
114	Rh054_05315	<<<	545	922302	922683	<<<	381	220	<<<	Rh054_05320
115	Rh054_05325	<<<	1172	926034	926196	<<<	162	561	<<<	Rh054_05345
116	Rh054_05375	>>>	5	932159	932504	>>>	345	443	<<<	Rh054_05390
117	Rh054_05395	>>>	932	934570	934843	>>>	273	231	>>>	Rh054_05400
118	Rh054_05400	>>>	663	935960	936280	<<<	320	312	<<<	Rh054_t07526
119	Rh054_05400	>>>	605	935902	936085	>>>	183	507	<<<	Rh054_t07526
120	Rh054_05605	>>>	215	959075	959166	<<<	91	42	<<<	Rh054_05610
121	Rh054_05635	<<<	33	965890	966003	<<<	113	6	<<<	Rh054_05640
122	Rh054_05765	>>>	37	983662	983805	>>>	143	1407	<<<	Rh054_05785
123	Rh054_05765	>>>	140	983765	983881	<<<	116	1331	<<<	Rh054_05785
124	Rh054_05765	>>>	213	983838	983881	<<<	43	1331	<<<	Rh054_05785
125	Rh054_05820	>>>	-4	990973	991107	<<<	134	6	<<<	Rh054_05825
126	Rh054_05910	>>>	128	1005830	1006091	>>>	261	-49	<<<	Rh054_05915
127	Rh054_05977	<<<	330	1023376	1023443	<<<	67	0	<<<	Rh054_06005
128	Rh054_05977	<<<	235	1023281	1023426	>>>	145	17	<<<	Rh054_06005

129	Rh054_06005	<<<	46	1028443	1028624	<<<	181	187	<<<	Rh054_06010
130	Rh054_06035	>>>	1	1033129	1033287	>>>	158	-41	<<<	Rh054_06040
131	Rh054_06245	>>>	56	1064614	1064739	<<<	125	14	<<<	Rh054_06250
132	Rh054_06475	>>>	125	1099660	1099747	>>>	87	70	>>>	Rh054_06480
133	Rh054_06510	<<<	118	1105420	1105525	<<<	105	129	<<<	Rh054_06515
134	Rh054_06555	>>>	1	1113047	1113147	>>>	100	60	>>>	Rh054_06560
135	Rh054_06655	>>>	1	1141123	1141232	>>>	109	-12	>>>	Rh054_06660
136	Rh054_06700	<<<	57	1149749	1149812	<<<	63	0	<<<	Rh054_06705
137	Rh054_06745	<<<	64	1162121	1162212	<<<	91	1222	<<<	Rh054_06750
138	Rh054_06830	>>>	106	1174015	1174110	>>>	95	1067	<<<	Rh054_06845
139	Rh054_06830	>>>	106	1174015	1174067	>>>	52	1110	<<<	Rh054_06845
140	Rh054_06830	>>>	96	1174005	1174442	<<<	437	735	<<<	Rh054_06845
141	Rh054_06960	>>>	40	1198462	1198595	>>>	133	223	<<<	Rh054_06965
142	Rh054_06960	>>>	40	1198462	1198559	>>>	97	259	<<<	Rh054_06965
143	Rh054_06960	>>>	40	1198462	1198805	>>>	343	13	<<<	Rh054_06965
144	Rh054_06960	>>>	331	1198753	1198818	<<<	65	0	<<<	Rh054_06965
145	Rh054_07075	>>>	62	1216981	1217142	>>>	161	1392	>>>	Rh054_07090
146	Rh054_07175	<<<	157	1234390	1234466	<<<	76	1	<<<	Rh054_07180
147	Rh054_07450	>>>	1	1274145	1274215	>>>	70	1769	<<<	Rh054_07465
148	Rh054_07450	>>>	1	1274145	1274197	>>>	52	1787	<<<	Rh054_07465
149	Rh054_07450	>>>	390	1274534	1274889	>>>	355	1095	<<<	Rh054_07465
150	Rh054_07450	>>>	1265	1275409	1275629	<<<	220	355	<<<	Rh054_07465
151	Rh054_07070	>>>	137	1215633	1215673	<<<	40	0	>>>	Rh054_07075
152	Rh054_00755	>>>	47	129967	130061	<<<	94	14	>>>	Rh054_00760
153	Rh054_00830	>>>	1	139818	139910	>>>	92	23	>>>	Rh054_00835
154	Rh054_01095	>>>	99	199722	199769	<<<	47	95	<<<	Rh054_01100

155	Rh054_03005	>>>	87	525351	525394	<<<	43	33	>>>	Rh054_03010
157	Rh054_t07508	>>>	-38	698183	698436	<<<	253	185	>>>	Rh054_04000
158	Rh054_04475	<<<	16	777708	777814	<<<	106	61	<<<	Rh054_04480
159	Rh054_04995	>>>	2	863221	863634	>>>	413	-41	>>>	Rh054_05010

Rickettsia japonica SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RJP_0009	<<<	3473	12368	12450	>>>	82	141	<<<	RJP_0010
2	RJP_0071	>>>	585	98307	98494	>>>	187	432	<<<	RJP_0072
3	RJP_0291	<<<	430	372082	372344	>>>	262	754	<<<	RJP_0292
4	RJP_0349	<<<	1	450600	450657	>>>	57	49	<<<	RJP_0350
5	RJP_0413	>>>	402	528332	528537	>>>	205	766	>>>	RJP_0414
6	RJP_0458	<<<	16	589113	589173	>>>	60	138	>>>	RJP_0459
7	RJP_0721	<<<	751	925236	925432	>>>	196	569	>>>	RJP_0722
8	RJP_0728	<<<	15	933403	933521	>>>	118	110	<<<	RJP_0729
9	RJP_0818	<<<	1230	1039414	1039494	>>>	80	183	<<<	RJP_0819
10	RJP_0994	>>>	242	1261706	1261784	<<<	78	27	>>>	RJP_0995
11	RJP_0986	>>>	-24	1255326	1255561	<<<	235	133	<<<	RJP_0987
12	RJP_0616	>>>	1628	789792	789942	<<<	150	1576	>>>	RJP_0617
13	RJP_0561	>>>	688	717968	718286	<<<	318	3105	>>>	RJP_0562
14	RJP_0524	>>>	1099	680216	680304	<<<	88	542	>>>	RJP_0525
15	RJP_0517	<<<	583	671652	671740	<<<	88	247	>>>	RJP_0519
16	RJP_0452	>>>	491	583867	584288	<<<	421	143	<<<	RJP_0453
17	RJP_0443	>>>	1860	566944	567170	<<<	226	3785	<<<	RJP_0444
18	RJP_0328	>>>	1124	422715	422761	<<<	46	2331	<<<	RJP_0329
19	RJP_0295	<<<	1918	378439	378503	<<<	64	1105	>>>	RJP_0296
20	RJP_0180	>>>	732	236330	236561	<<<	231	685	<<<	RJP_0181
21	RJP_0107	>>>	84	137769	137874	<<<	105	25	>>>	RJP_0108
22	RJP_0015	>>>	1	26094	26209	>>>	115	-4	>>>	RJP_0016
23	RJP_0022	<<<	18	31971	32061	<<<	90	0	<<<	RJP_0023
24	RJP_0031	>>>	55	45114	45451	<<<	337	3223	<<<	RJP_0032

25	RJP_0031	>>>	2370	47429	47556	>>>	127	1118	<<<	RJP_0032
26	RJP_0031	>>>	3520	48579	48627	<<<	48	47	<<<	RJP_0032
27	RJP_0034	>>>	116	56778	56839	<<<	61	43	<<<	RJP_0035
28	RJP_0043	>>>	357	68541	68586	>>>	45	147	<<<	RJP_0044
29	RJP_0046	>>>	34	72155	72238	>>>	83	64	>>>	RJP_0047
30	RJP_0066	>>>	1	90750	90881	>>>	131	-10	<<<	RJP_0067
31	RJP_0070	>>>	9	95327	95467	>>>	140	40	>>>	RJP_0071
32	RJP_0071	>>>	10	97732	97856	>>>	124	1070	<<<	RJP_0072
33	RJP_0071	>>>	850	98572	98917	<<<	345	9	<<<	RJP_0072
34	RJP_0076	>>>	16	103376	103524	>>>	148	153	<<<	RJP_0077
35	RJP_0087	>>>	60	119190	119246	>>>	56	70	>>>	RJP_0088
36	RJP_0088	>>>	123	120370	120428	<<<	58	1326	<<<	RJP_0089
37	RJP_0088	>>>	155	120402	120719	<<<	317	1035	<<<	RJP_0089
38	RJP_0088	>>>	863	121110	121510	>>>	400	244	<<<	RJP_0089
39	RJP_0108	>>>	1	140319	140441	>>>	122	-7	>>>	RJP_0109
40	RJP_0128	>>>	8	172421	172555	>>>	134	33	>>>	RJP_0129
41	RJP_0138	>>>	1	181506	181676	>>>	170	98	>>>	RJP_0139
42	RJP_0143	>>>	1	189153	189227	>>>	74	165	>>>	RJP_0144
43	RJP_0144	>>>	258	193770	193910	>>>	140	-1	<<<	RJP_0145
44	RJP_0149	>>>	105	200298	200338	<<<	40	95	<<<	RJP_0150
45	RJP_0163	>>>	-4	213934	214219	<<<	285	86	<<<	RJP_0164
46	RJP_0179	>>>	1389	233084	233232	>>>	148	1162	>>>	RJP_0180
47	RJP_0185	<<<	57	240874	240980	<<<	106	0	<<<	RJP_0186
48	RJP_0187	<<<	40	245435	245557	<<<	122	13	<<<	RJP_0188
49	RJP_0193	>>>	20	253782	253901	<<<	119	0	<<<	RJP_0194
50	RJP_0199	>>>	1	258465	258551	>>>	86	430	<<<	RJP_0200

51	RJP_0199	>>>	25	258489	258623	<<<	134	358	<<<	RJP_0200
52	RJP_0199	>>>	217	258681	258764	>>>	83	217	<<<	RJP_0200
53	RJP_0199	>>>	242	258706	258783	<<<	77	198	<<<	RJP_0200
54	RJP_0201	>>>	115	260546	260597	>>>	51	1673	<<<	RJP_0202
55	RJP_0201	>>>	108	260539	260641	<<<	102	1629	<<<	RJP_0202
56	RJP_0202	<<<	2852	267541	267585	<<<	44	2268	<<<	RJP_0203
57	RJP_0202	<<<	3883	268572	268855	<<<	283	998	<<<	RJP_0203
58	RJP_0236	>>>	2481	311791	312185	>>>	394	48	<<<	RJP_0237
59	RJP_0263	>>>	14	341702	341793	>>>	91	31	>>>	RJP_0264
60	RJP_0266	>>>	25	344979	345097	<<<	118	0	<<<	RJP_0267
61	RJP_0272	>>>	10	352167	352296	>>>	129	107	>>>	RJP_0273
62	RJP_0288	>>>	15	368651	368762	>>>	111	383	>>>	RJP_0289
63	RJP_0299	<<<	81	386870	386970	<<<	100	1	<<<	RJP_0300
64	RJP_0311	>>>	1	398685	398796	>>>	111	39	>>>	RJP_0312
65	RJP_0328	>>>	117	421708	421828	>>>	120	3264	<<<	RJP_0329
66	RJP_0348	>>>	153	446684	446739	<<<	55	2515	<<<	RJP_0349
67	RJP_0348	>>>	2612	449143	449253	<<<	110	1	<<<	RJP_0349
68	RJP_0348	>>>	2577	449108	449201	>>>	93	53	<<<	RJP_0349
69	RJP_0355	>>>	8	458506	458675	>>>	169	425	>>>	RJP_0356
70	RJP_0366	<<<	958	472459	472668	<<<	209	0	<<<	RJP_0367
71	RJP_0376	>>>	499	483709	483854	>>>	145	107	<<<	RJP_0377
72	RJP_0400	>>>	1	508032	508130	>>>	98	39	>>>	RJP_0401
73	RJP_0404	>>>	21	513696	513841	>>>	145	2779	<<<	RJP_0405
74	RJP_0404	>>>	104	513779	513894	<<<	115	2726	<<<	RJP_0405
75	RJP_0405	<<<	153	517719	517996	<<<	277	1033	<<<	RJP_0406
76	RJP_0413	>>>	9	527939	528166	>>>	227	1137	>>>	RJP_0414

77	RJP_0438	>>>	41	558470	558519	>>>	49	84	>>>	RJP_0439
78	RJP_0438	>>>	41	558470	558563	>>>	93	40	>>>	RJP_0439
79	RJP_0439	>>>	974	560202	560330	>>>	128	692	>>>	RJP_0440
80	RJP_0439	>>>	1663	560891	560923	>>>	32	99	>>>	RJP_0440
81	RJP_0439	>>>	1680	560908	560950	>>>	42	72	>>>	RJP_0440
82	RJP_0443	>>>	6	565090	565184	>>>	94	5771	<<<	RJP_0444
83	RJP_0443	>>>	2310	567394	567441	<<<	47	3514	<<<	RJP_0444
84	RJP_0452	>>>	150	583526	583588	>>>	62	843	<<<	RJP_0453
85	RJP_0452	>>>	849	584225	584283	<<<	58	148	<<<	RJP_0453
86	RJP_0456	<<<	31	586888	587001	<<<	113	0	<<<	RJP_0457
87	RJP_0466	<<<	693	596422	596508	<<<	86	0	<<<	RJP_0467
88	RJP_0469	<<<	104	600141	600298	<<<	157	433	<<<	RJP_0470
89	RJP_0473	>>>	459	603401	603536	>>>	135	318	<<<	RJP_0474
90	RJP_0475	<<<	244	605008	605091	<<<	83	0	<<<	RJP_0476
91	RJP_0484	>>>	15	615120	615173	>>>	53	462	>>>	RJP_0485
92	RJP_0487	>>>	1	620907	621044	>>>	137	37	<<<	RJP_0488
93	RJP_0489	>>>	4	624540	624848	>>>	308	2621	<<<	RJP_0490
94	RJP_0489	>>>	4	624540	624818	>>>	278	2651	<<<	RJP_0490
95	RJP_0491	>>>	1	628258	628373	>>>	115	-1	>>>	RJP_0492
96	RJP_0495	>>>	2535	635325	635440	>>>	115	29	>>>	RJP_0496
97	RJP_0501	<<<	37	648895	649019	<<<	124	0	<<<	RJP_0502
98	RJP_0516	<<<	91	668258	668366	<<<	108	53	<<<	RJP_0517
99	RJP_0521	>>>	14	674150	674251	>>>	101	67	>>>	RJP_0522
100	RJP_0542	>>>	1	696299	696395	>>>	96	170	>>>	RJP_0543
101	RJP_0548	<<<	116	702773	702906	<<<	133	0	<<<	RJP_0549
102	RJP_0556	>>>	12	712731	712931	>>>	200	54	>>>	RJP_0557

103	RJP_0557	>>>	209	713738	713886	>>>	148	-22	<<<	RJP_0558
104	RJP_0568	>>>	4	729146	729291	>>>	145	84	<<<	RJP_0569
105	RJP_0587	<<<	140	753007	753127	<<<	120	30	<<<	RJP_0588
106	RJP_0595	>>>	153	763778	763885	>>>	107	64	>>>	RJP_0596
107	RJP_0596	>>>	1	765436	765555	>>>	119	17	<<<	RJP_0597
108	RJP_0612	<<<	54	784567	784670	<<<	103	0	<<<	RJP_0613
109	RJP_0615	>>>	2	786830	787162	>>>	332	924	>>>	RJP_0616
110	RJP_0623	>>>	111	807380	807486	<<<	106	3	<<<	RJP_0624
111	RJP_0624	<<<	277	808727	808807	<<<	80	316	<<<	RJP_0625
112	RJP_0643	>>>	1	830892	830987	>>>	95	77	<<<	RJP_0644
113	RJP_0663	<<<	1659	850616	850653	<<<	37	267	<<<	RJP_0664
114	RJP_0676	>>>	223	863134	863553	>>>	419	520	>>>	RJP_0677
115	RJP_0677	>>>	183	864650	864917	>>>	267	222	>>>	RJP_0678
116	RJP_0680	<<<	61	867874	868100	<<<	226	10	<<<	RJP_0681
117	RJP_0680	<<<	192	868005	868100	<<<	95	10	<<<	RJP_0681
118	RJP_0683	>>>	56	870259	870333	<<<	74	10	<<<	RJP_0684
119	RJP_0689	<<<	76	878581	878659	<<<	78	1317	<<<	RJP_0690
120	RJP_0689	<<<	200	878705	878752	<<<	47	1224	<<<	RJP_0690
121	RJP_0689	<<<	229	878734	879193	<<<	459	783	<<<	RJP_0690
122	RJP_0694	>>>	1	885830	885914	>>>	84	279	>>>	RJP_0695
123	RJP_0697	>>>	1734	895600	895813	>>>	213	-22	<<<	RJP_0698
124	RJP_0699	>>>	48	897663	897781	<<<	118	247	<<<	RJP_0700
125	RJP_0707	>>>	242	904264	904423	>>>	159	482	>>>	RJP_0708
126	RJP_0711	>>>	1057	909717	909995	<<<	278	648	<<<	RJP_0712
127	RJP_0716	>>>	24	917820	917963	>>>	143	40	>>>	RJP_0717
128	RJP_0725	<<<	54	931487	931618	<<<	131	16	<<<	RJP_0726

129	RJP_0733	<<<	356	936775	936995	>>>	220	2854	<<<	RJP_0734
130	RJP_0733	<<<	1653	938072	938200	<<<	128	1649	<<<	RJP_0734
131	RJP_0733	<<<	2444	938863	939014	>>>	151	835	<<<	RJP_0734
132	RJP_0762	<<<	17	953817	954037	<<<	220	2	<<<	RJP_0763
133	RJP_0770	>>>	1454	962294	962385	<<<	91	42	<<<	RJP_0771
134	RJP_0781	>>>	1490	978969	979094	>>>	125	2017	<<<	RJP_0782
135	RJP_0786	>>>	37	986869	987012	>>>	143	2385	<<<	RJP_0787
136	RJP_0786	>>>	140	986972	987071	<<<	99	2326	<<<	RJP_0787
137	RJP_0786	>>>	213	987045	987133	<<<	88	2264	<<<	RJP_0787
138	RJP_0790	>>>	-4	994179	994313	<<<	134	6	<<<	RJP_0791
139	RJP_0795	>>>	3725	1006253	1006480	>>>	227	1894	<<<	RJP_0796
140	RJP_0795	>>>	5635	1008163	1008423	>>>	260	-49	<<<	RJP_0796
141	RJP_0801	>>>	7	1013788	1013919	>>>	131	163	>>>	RJP_0802
142	RJP_0808	>>>	1180	1024886	1024965	>>>	79	723	<<<	RJP_0809
143	RJP_0808	>>>	1820	1025526	1025674	>>>	148	14	<<<	RJP_0809
144	RJP_0808	>>>	1914	1025620	1025688	<<<	68	0	<<<	RJP_0809
145	RJP_0809	<<<	45	1030705	1030888	<<<	183	187	<<<	RJP_0810
146	RJP_0815	>>>	1	1035395	1035553	>>>	158	-41	<<<	RJP_0816
147	RJP_0829	>>>	1	1051940	1052069	>>>	129	-1	>>>	RJP_0830
148	RJP_0874	>>>	125	1102191	1102278	>>>	87	70	>>>	RJP_0875
149	RJP_0877	>>>	49	1105331	1105544	<<<	213	0	<<<	RJP_0878
150	RJP_0877	>>>	1	1105283	1105586	>>>	303	-42	<<<	RJP_0878
151	RJP_0880	<<<	118	1107950	1108055	<<<	105	129	<<<	RJP_0881
152	RJP_0890	>>>	1785	1120159	1120311	<<<	152	1984	<<<	RJP_0891
153	RJP_0890	>>>	2093	1120467	1120560	<<<	93	1735	<<<	RJP_0891
154	RJP_0900	>>>	1616	1143499	1143736	>>>	237	680	>>>	RJP_0901

155	RJP_0907	<<<	57	1152282	1152345	<<<	63	0	<<<	RJP_0908
156	RJP_0923	>>>	106	1176565	1176660	>>>	95	4356	>>>	RJP_0924
157	RJP_0923	>>>	106	1176565	1176617	>>>	52	4399	>>>	RJP_0924
158	RJP_0923	>>>	1947	1178406	1178593	>>>	187	2423	>>>	RJP_0924
159	RJP_0937	>>>	6	1202178	1202346	>>>	168	223	<<<	RJP_0938
160	RJP_0937	>>>	6	1202178	1202310	>>>	132	259	<<<	RJP_0938
161	RJP_0937	>>>	6	1202178	1202556	>>>	378	13	<<<	RJP_0938
162	RJP_0937	>>>	332	1202504	1202569	<<<	65	0	<<<	RJP_0938
163	RJP_0957	>>>	166	1221038	1221119	>>>	81	609	>>>	RJP_0958
164	RJP_0971	<<<	156	1238283	1238359	<<<	76	1	<<<	RJP_0972
165	RJP_1007	>>>	1	1278342	1278412	>>>	70	1793	<<<	RJP_1008
166	RJP_1007	>>>	390	1278731	1279085	>>>	354	1120	<<<	RJP_1008
167	RJP_1007	>>>	1264	1279605	1279826	<<<	221	379	<<<	RJP_1008
168	RJP_0845	<<<	1	1072894	1073034	>>>	140	-28	<<<	RJP_0846
169	RJP_0084	>>>	72	110052	110586	<<<	534	330	<<<	RJP_0085
170	RJP_0097	>>>	47	130596	130690	<<<	94	14	>>>	RJP_0098
171	RJP_0108	>>>	1	140319	140411	>>>	92	23	>>>	RJP_0109
172	RJP_0179	>>>	224	231919	231958	>>>	39	2436	>>>	RJP_0180
173	RJP_0288	>>>	15	368651	368725	>>>	74	420	>>>	RJP_0289
174	RJP_0289	>>>	144	370046	370161	<<<	115	530	>>>	RJP_0290
175	RJP_0545	>>>	-38	699235	699598	<<<	363	490	>>>	RJP_0546
176	RJP_0610	<<<	16	781785	781891	<<<	106	1442	>>>	RJP_0611
177	RJP_0681	<<<	1	868982	869035	>>>	53	82	>>>	RJP_0682
178	RJP_0719	>>>	336	922740	923146	<<<	406	0	<<<	RJP_0720

Rickettsia massiliae SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RMA_0010	<<<	1	11500	11924	>>>	424	104	<<<	RMA_RNA01
2	RMA_0462	<<<	1	458964	459021	>>>	57	49	<<<	RMA_0463
3	RMA_0608	<<<	1	599068	599128	>>>	60	138	>>>	RMA_0609
4	RMA_0832	<<<	54	825682	825799	>>>	117	1049	<<<	RMA_0833
5	RMA_1368	>>>	255	1336501	1336795	>>>	294	-24	<<<	RMA_1369
6	RMA_1108	>>>	85	1086295	1086449	<<<	154	23	>>>	RMA_1109
7	RMA_0955	<<<	108	952091	952175	<<<	84	0	<<<	RMA_0956
8	RMA_0926	>>>	105	924338	924435	<<<	97	22	>>>	RMA_0927
9	RMA_0683	<<<	582	680358	680449	<<<	91	251	>>>	RMA_0684
10	RMA_0602	>>>	519	594035	594115	<<<	80	273	<<<	RMA_0603
11	RMA_0375	<<<	64	381876	382034	<<<	158	132	<<<	RMA_0376
12	RMA_0356	<<<	1081	364762	364836	<<<	74	1	>>>	RMA_0357
13	RMA_0304	>>>	565	310473	310519	<<<	46	0	>>>	RMA_0305
14	RMA_0265	>>>	-53	271024	271257	<<<	233	4	<<<	RMA_0266
15	RMA_0009	>>>	106	10563	10916	<<<	353	300	<<<	RMA_0010
16	RMA_0063	>>>	116	55706	55767	<<<	61	61	<<<	RMA_0064
17	RMA_0079	>>>	36	70682	70765	>>>	83	64	>>>	RMA_0080
18	RMA_0086	>>>	4	75990	76065	>>>	75	115	<<<	RMA_0087
19	RMA_0105	>>>	1	96402	96526	>>>	124	1090	<<<	RMA_0106
20	RMA_0105	>>>	860	97261	97539	<<<	278	77	<<<	RMA_0106
21	RMA_0110	>>>	16	102059	102191	>>>	132	175	<<<	RMA_0111
22	RMA_0117	>>>	76	108745	108878	<<<	133	402	>>>	RMA_0118
23	RMA_0121	>>>	60	117852	117908	>>>	56	70	>>>	RMA_0122
24	RMA_0122	>>>	123	119032	119090	<<<	58	386	>>>	RMA_0123

25	RMA_0122	>>>	155	119064	119110	<<<	46	366	>>>	RMA_0123
26	RMA_0145	>>>	1	138967	139089	>>>	122	-7	>>>	RMA_0146
27	RMA_0148	>>>	1	147374	147541	>>>	167	5	>>>	RMA_0149
28	RMA_0165	>>>	2	163747	163878	>>>	131	-23	>>>	RMA_0166
29	RMA_0184	>>>	64	180325	180433	>>>	108	99	>>>	RMA_0185
30	RMA_0188	>>>	54	182942	183087	>>>	145	648	>>>	RMA_0189
31	RMA_0188	>>>	580	183468	183515	>>>	47	220	>>>	RMA_0189
32	RMA_0189	>>>	25	187883	187933	>>>	50	177	>>>	RMA_0190
33	RMA_0190	>>>	140	192376	192639	>>>	263	-1	<<<	RMA_0191
34	RMA_0208	>>>	-4	212795	212939	<<<	144	486	<<<	RMA_0209
35	RMA_0231	>>>	899	235498	235568	>>>	70	310	<<<	RMA_0232
36	RMA_0238	<<<	54	240588	240694	<<<	106	0	<<<	RMA_0239
37	RMA_0238	<<<	75	240609	240694	<<<	85	0	<<<	RMA_0239
38	RMA_0255	>>>	1	259547	259632	>>>	85	346	<<<	RMA_0256
39	RMA_0255	>>>	24	259570	259650	<<<	80	328	<<<	RMA_0256
40	RMA_0257	>>>	107	261619	261720	<<<	101	1087	<<<	RMA_0258
41	RMA_0257	>>>	114	261626	261677	>>>	51	1130	<<<	RMA_0258
42	RMA_0263	>>>	719	268953	268985	<<<	32	1269	>>>	RMA_0264
43	RMA_0265	>>>	59	271136	271257	<<<	121	4	<<<	RMA_0266
44	RMA_0276	>>>	25	284621	284822	>>>	201	646	>>>	RMA_0277
45	RMA_0280	>>>	218	288924	289125	<<<	201	427	>>>	RMA_RNA07
46	RMA_0307	>>>	758	313956	314323	>>>	367	62	<<<	RMA_0308
47	RMA_0326	>>>	23	332267	332527	>>>	260	-24	<<<	RMA_0327
48	RMA_0340	>>>	12	344503	344609	>>>	106	123	>>>	RMA_0341
49	RMA_0340	>>>	12	344503	344651	>>>	148	81	>>>	RMA_0341
50	RMA_0343	>>>	28	347891	348007	<<<	116	0	<<<	RMA_0344

51	RMA_0349	>>>	10	355098	355149	>>>	51	180	>>>	RMA_0350
52	RMA_0349	>>>	36	355124	355225	>>>	101	104	>>>	RMA_0350
53	RMA_0364	>>>	7	371426	371545	>>>	119	-10	>>>	RMA_0365
54	RMA_0388	<<<	73	392622	392695	<<<	73	28	<<<	RMA_0389
55	RMA_0416	>>>	100	420698	420747	>>>	49	4	<<<	RMA_0417
56	RMA_0416	>>>	75	420673	420740	<<<	67	11	<<<	RMA_0417
57	RMA_0416	>>>	17	420615	420717	>>>	102	34	<<<	RMA_0417
58	RMA_0422	>>>	102	427774	427909	>>>	135	121	>>>	RMA_0423
59	RMA_0422	>>>	197	427869	428007	<<<	138	23	>>>	RMA_0423
60	RMA_0440	<<<	45	444109	444195	<<<	86	0	<<<	RMA_0441
61	RMA_0483	<<<	1223	480754	480904	<<<	150	59	<<<	RMA_0484
62	RMA_0487	>>>	514	484678	484711	<<<	33	299	<<<	RMA_0488
63	RMA_0511	>>>	28	504548	504618	>>>	70	89	<<<	RMA_0512
64	RMA_0515	>>>	388	508966	509247	>>>	281	-17	>>>	RMA_0516
65	RMA_0538	<<<	146	527254	527599	<<<	345	0	>>>	RMA_0539
66	RMA_0546	>>>	4	537416	537643	>>>	227	-28	>>>	RMA_0547
67	RMA_0549	>>>	46	539650	539702	>>>	52	1071	>>>	RMA_0550
68	RMA_0563	>>>	8	554209	554344	>>>	135	66	>>>	RMA_0564
69	RMA_0600	>>>	216	590217	590305	>>>	88	61	>>>	RMA_0601
70	RMA_0602	>>>	12	593528	593724	>>>	196	664	<<<	RMA_0603
71	RMA_0606	<<<	31	596840	596953	<<<	113	0	<<<	RMA_0607
72	RMA_0611	>>>	1	601666	601795	>>>	129	-23	>>>	RMA_0612
73	RMA_0617	<<<	687	606401	606487	<<<	86	0	<<<	RMA_0618
74	RMA_0619	<<<	730	609978	610133	<<<	155	440	<<<	RMA_0620
75	RMA_0630	>>>	34	621400	621532	>>>	132	25	>>>	RMA_0631
76	RMA_0633	>>>	1	624698	624765	>>>	67	88	>>>	RMA_0634

77	RMA_0635	>>>	12	626055	626189	>>>	134	37	>>>	RMA_0636
78	RMA_0639	>>>	1	634635	634939	>>>	304	130	>>>	RMA_0640
79	RMA_0639	>>>	1	634635	634909	>>>	274	160	>>>	RMA_0640
80	RMA_0658	>>>	16	651673	651926	>>>	253	121	>>>	RMA_0659
81	RMA_0664	<<<	45	656771	656895	<<<	124	0	<<<	RMA_0665
82	RMA_0664	<<<	74	656800	656895	<<<	95	0	<<<	RMA_0665
83	RMA_0682	<<<	85	676963	677113	<<<	150	10	<<<	RMA_0683
84	RMA_0686	>>>	45	682896	682966	>>>	70	60	>>>	RMA_0687
85	RMA_RNA18	<<<	116	774417	774547	<<<	130	0	<<<	RMA_0774
86	RMA_0786	>>>	59	784215	784368	>>>	153	34	>>>	RMA_0787
87	RMA_0806	>>>	4	800216	800362	>>>	146	86	<<<	RMA_0807
88	RMA_0839	>>>	131	835548	835677	>>>	129	117	>>>	RMA_0840
89	RMA_0840	>>>	1	837284	837413	>>>	129	90	<<<	RMA_0841
90	RMA_0844	<<<	1	838536	838578	>>>	42	415	>>>	RMA_0845
91	RMA_0859	<<<	36	856200	856303	<<<	103	0	<<<	RMA_0860
92	RMA_0889	>>>	107	884771	884877	<<<	106	3	<<<	RMA_0890
93	RMA_0890	<<<	230	886116	886516	<<<	400	0	<<<	RMA_0891
94	RMA_0908	<<<	179	905036	905176	<<<	140	0	<<<	RMA_0909
95	RMA_0930	>>>	18	926248	926401	>>>	153	31	>>>	RMA_0931
96	RMA_0940	>>>	5	935139	935315	>>>	176	-28	<<<	RMA_0941
97	RMA_0943	<<<	60	939696	939816	<<<	120	22	<<<	RMA_0944
98	RMA_0949	>>>	173	947651	947943	>>>	292	-21	<<<	RMA_0950
99	RMA_0966	>>>	243	959286	959316	>>>	30	185	<<<	RMA_0967
100	RMA_0966	>>>	243	959286	959441	>>>	155	60	<<<	RMA_0967
101	RMA_0967	<<<	151	961180	961264	<<<	84	0	<<<	RMA_0968
102	RMA_0972	>>>	57	965573	965606	>>>	33	898	<<<	RMA_0973

103	RMA_0984	>>>	561	980464	980561	>>>	97	670	<<<	RMA_0985
104	RMA_0984	>>>	578	980481	980669	<<<	188	562	<<<	RMA_0985
105	RMA_0984	>>>	561	980464	980525	>>>	61	706	<<<	RMA_0985
106	RMA_0984	>>>	657	980560	980675	>>>	115	556	<<<	RMA_0985
107	RMA_0984	>>>	680	980583	980669	<<<	86	562	<<<	RMA_0985
108	RMA_0988	>>>	1	989476	989554	>>>	78	171	<<<	RMA_0989
109	RMA_0988	>>>	1	989476	989573	>>>	97	152	<<<	RMA_0989
110	RMA_0991	>>>	6	991722	991905	>>>	183	26	>>>	RMA_0992
111	RMA_0996	<<<	371	997504	997765	<<<	261	326	<<<	RMA_0997
112	RMA_1001	<<<	96	1002361	1002491	<<<	130	2	<<<	RMA_1002
113	RMA_1005	<<<	605	1004886	1004947	<<<	61	0	<<<	RMA_1006
114	RMA_1012	>>>	234	1009825	1009877	>>>	52	190	<<<	RMA_1013
115	RMA_1013	<<<	207	1010605	1010757	>>>	152	829	<<<	RMA_RNA27
116	RMA_1013	<<<	549	1010947	1011021	<<<	74	565	<<<	RMA_RNA27
117	RMA_1013	<<<	207	1010605	1011072	>>>	467	514	<<<	RMA_RNA27
118	RMA_1013	<<<	958	1011356	1011394	>>>	38	192	<<<	RMA_RNA27
119	RMA_1050	>>>	226	1034013	1034103	<<<	90	47	<<<	RMA_1051
120	RMA_1050	>>>	189	1033976	1034063	>>>	87	87	<<<	RMA_1051
121	RMA_1053	<<<	408	1038240	1038330	<<<	90	3	<<<	RMA_1054
122	RMA_1063	>>>	195	1048214	1048299	>>>	85	75	>>>	RMA_1064
123	RMA_1083	>>>	5	1066099	1066252	<<<	153	0	<<<	RMA_1084
124	RMA_1083	>>>	10	1066104	1066159	>>>	55	93	<<<	RMA_1084
125	RMA_1103	>>>	69	1080638	1080760	>>>	122	89	<<<	RMA_1104
126	RMA_1108	>>>	7	1086217	1086306	>>>	89	166	>>>	RMA_1109
127	RMA_1108	>>>	7	1086217	1086348	>>>	131	124	>>>	RMA_1109
128	RMA_1112	>>>	32	1090736	1090805	>>>	69	109	>>>	RMA_1113

129	RMA_1117	<<<	156	1098472	1098563	>>>	91	14	<<<	RMA_1118
130	RMA_1117	<<<	193	1098509	1098577	<<<	68	0	<<<	RMA_1118
131	RMA_1118	<<<	42	1103588	1103771	<<<	183	78	<<<	RMA_1119
132	RMA_1166	>>>	57	1138258	1138398	<<<	140	326	<<<	RMA_1167
134	RMA_1183	>>>	1	1150737	1150883	>>>	146	12	<<<	RMA_1184
135	RMA_1183	>>>	1	1150737	1150856	>>>	119	39	<<<	RMA_1184
136	RMA_1199	>>>	103	1167251	1167507	<<<	256	457	<<<	RMA_1200
137	RMA_1206	>>>	125	1173241	1173328	>>>	87	69	>>>	RMA_1207
138	RMA_1208	>>>	48	1176379	1176593	<<<	214	0	<<<	RMA_1209
139	RMA_1208	>>>	1	1176332	1176635	>>>	303	-42	<<<	RMA_1209
140	RMA_1224	>>>	132	1191069	1191219	<<<	150	0	>>>	RMA_1225
141	RMA_1306	>>>	18	1279233	1279391	>>>	158	223	<<<	RMA_1307
143	RMA_1306	>>>	333	1279548	1279614	<<<	66	0	<<<	RMA_1307
144	RMA_1318	<<<	96	1289480	1289610	<<<	130	348	<<<	RMA_1319
145	RMA_1324	>>>	1	1296233	1296302	>>>	69	66	>>>	RMA_1325
146	RMA_1326	>>>	63	1299736	1299848	>>>	112	553	>>>	RMA_1327
147	RMA_1347	>>>	1	1321955	1322065	>>>	110	65	>>>	RMA_RNA37
148	RMA_1382	>>>	3	1353403	1353509	>>>	106	183	>>>	RMA_1383
149	RMA_1386	<<<	512	1357398	1357617	<<<	219	400	<<<	RMA_1387
150	RMA_1386	<<<	535	1357421	1357482	>>>	61	535	<<<	RMA_1387
151	RMA_1386	<<<	542	1357428	1357617	<<<	189	400	<<<	RMA_1387
152	RMA_0122	>>>	37	118946	119035	>>>	89	441	>>>	RMA_0123
153	RMA_0223	>>>	6	230662	230915	>>>	253	122	<<<	RMA_0224
154	RMA_0249	<<<	141	253925	254051	<<<	126	814	<<<	RMA_0250
155	RMA_0364	>>>	7	371426	371508	>>>	82	27	>>>	RMA_0365
156	RMA_0545	>>>	87	536053	536097	<<<	44	0	>>>	RMA_0546

157	RMA_RNA17	>>>	-38	767051	767299	<<<	248	47	>>>	RMA_0761
158	RMA_0855	<<<	7	853514	853621	<<<	107	228	<<<	RMA_0856
159	RMA_0969	<<<	92	962825	963035	>>>	210	37	<<<	RMA_0970

Rickettsia peacockii SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RPR_00205	<<<	246	31463	31552	>>>	89	594	>>>	RPR_00210
2	RPR_00380	<<<	1068	69661	70206	>>>	545	1015	<<<	RPR_00420
3	RPR_02035	<<<	532	332464	332518	>>>	54	1110	<<<	RPR_02050
4	RPR_02300	>>>	883	386381	386643	>>>	262	754	<<<	RPR_02325
5	RPR_02575	>>>	116	435561	435707	>>>	146	436	<<<	RPR_02585
6	RPR_02850	>>>	2528	494355	494484	>>>	129	566	>>>	RPR_02885
7	RPR_04730	>>>	3131	809690	809954	>>>	264	4685	<<<	RPR_04800
8	RPR_04970	<<<	1045	862212	862452	>>>	240	4	<<<	RPR_04980
9	RPR_06535	>>>	273	1097311	1097524	>>>	213	1255	>>>	RPR_06550
10	RPR_07615	>>>	233	1267419	1267497	<<<	78	27	>>>	RPR_07620
11	RPR_07555	>>>	-24	1261440	1261672	<<<	232	668	<<<	RPR_07575
12	RPR_07080	>>>	49	1183937	1183995	<<<	58	0	>>>	RPR_07085
13	RPR_06420	>>>	238	1082277	1082339	<<<	62	1641	>>>	RPR_06445
14	RPR_05550	<<<	138	946667	946729	<<<	62	14	>>>	RPR_05555
15	RPR_05135	>>>	4759	892589	892765	<<<	176	369	>>>	RPR_05175
16	RPR_03760	>>>	1514	631176	631488	<<<	312	18	<<<	RPR_03775
17	RPR_03560	>>>	110	599030	599127	<<<	97	36	>>>	RPR_03565
18	RPR_03380	>>>	171	569469	569699	<<<	230	117	>>>	RPR_03390
19	RPR_02850	>>>	2062	493889	494356	<<<	467	694	>>>	RPR_02885
20	RPR_02820	>>>	84	474852	474957	<<<	105	25	>>>	RPR_02825
21	RPR_02725	>>>	522	465750	466222	<<<	472	244	>>>	RPR_02750
22	RPR_02285	>>>	440	384646	384763	<<<	117	230	>>>	RPR_02300
23	RPR_00350	<<<	388	64629	65098	<<<	469	2	<<<	RPR_00365
24	RPR_00060	<<<	512	13394	13761	<<<	367	152	<<<	RPR_00090

25	RPR_00050	>>>	139	10624	10678	<<<	54	449	<<<	RPR_00055
26	RPR_00105	<<<	59	18865	18931	<<<	66	32	<<<	RPR_00110
27	RPR_00145	<<<	1572	26454	26617	<<<	163	174	<<<	RPR_00170
28	RPR_00210	>>>	12	34796	34991	>>>	195	54	>>>	RPR_00215
29	RPR_00210	>>>	12	34796	34961	>>>	165	84	>>>	RPR_00215
30	RPR_00255	>>>	47	42990	43060	<<<	70	0	<<<	RPR_00260
31	RPR_00300	>>>	1	54993	55134	>>>	141	27	>>>	RPR_00310
32	RPR_00380	<<<	64	68657	68739	<<<	82	2482	<<<	RPR_00420
33	RPR_00455	<<<	25	77031	77121	<<<	90	0	<<<	RPR_00460
34	RPR_00495	>>>	54	84282	84401	<<<	119	260	<<<	RPR_00505
35	RPR_00535	<<<	101	90658	90688	<<<	30	35	<<<	RPR_00540
36	RPR_00560	<<<	58	94315	94457	<<<	142	0	<<<	RPR_00565
37	RPR_00570	<<<	1287	97616	97759	<<<	143	18	<<<	RPR_00585
38	RPR_00590	>>>	113	100428	100460	<<<	32	1731	<<<	RPR_00600
39	RPR_00590	>>>	417	100732	100871	<<<	139	1320	<<<	RPR_00600
40	RPR_00615	>>>	401	104814	104997	>>>	183	842	>>>	RPR_00625
41	RPR_00630	>>>	1	108613	108714	>>>	101	674	>>>	RPR_00645
42	RPR_00695	>>>	153	123396	123503	>>>	107	105	>>>	RPR_00705
43	RPR_00705	>>>	1	125095	125251	>>>	156	2	<<<	RPR_00710
44	RPR_00705	>>>	1	125095	125226	>>>	131	27	<<<	RPR_00710
45	RPR_00755	>>>	1	133092	133160	>>>	68	52	>>>	RPR_00760
46	RPR_00965	<<<	54	163146	163249	<<<	103	0	<<<	RPR_00970
47	RPR_01080	<<<	332	177498	177587	<<<	89	10	<<<	RPR_01090
48	RPR_01280	>>>	1	208312	208465	>>>	153	228	>>>	RPR_01285
49	RPR_01290	>>>	1836	213897	214044	>>>	147	129	<<<	RPR_01315
50	RPR_01290	>>>	1980	214041	214159	>>>	118	14	<<<	RPR_01315

51	RPR_01290	>>>	2044	214105	214173	<<<	68	0	<<<	RPR_01315
52	RPR_01345	>>>	1	223851	224009	>>>	158	-41	<<<	RPR_01350
53	RPR_01475	>>>	1	240651	240786	>>>	135	-11	>>>	RPR_01480
54	RPR_01735	>>>	72	279802	279879	>>>	77	2894	<<<	RPR_01775
55	RPR_01735	>>>	88	279818	280162	<<<	344	2611	<<<	RPR_01775
56	RPR_01735	>>>	72	279802	279922	>>>	120	2851	<<<	RPR_01775
57	RPR_01855	>>>	58	297983	298130	<<<	147	1255	<<<	RPR_01870
58	RPR_01855	>>>	1	297926	298023	>>>	97	1362	<<<	RPR_01870
59	RPR_01985	>>>	97	323761	323846	<<<	85	2171	<<<	RPR_02005
60	RPR_01985	>>>	6	323670	323801	>>>	131	2216	<<<	RPR_02005
61	RPR_02010	>>>	156	327707	327893	<<<	186	0	<<<	RPR_02015
62	RPR_02010	>>>	1	327552	327765	>>>	213	128	<<<	RPR_02015
63	RPR_02010	>>>	1	327552	327894	>>>	342	-1	<<<	RPR_02015
64	RPR_02075	>>>	15	341277	341422	>>>	145	15	>>>	RPR_02080
65	RPR_02085	>>>	1	345853	345972	>>>	119	43	>>>	RPR_02090
66	RPR_02145	>>>	12	355975	356082	>>>	107	123	>>>	RPR_02150
67	RPR_02160	>>>	25	359366	359482	<<<	116	0	<<<	RPR_02165
68	RPR_02190	>>>	39	366582	366682	>>>	100	152	>>>	RPR_02195
69	RPR_02245	>>>	19	376835	376967	>>>	132	25	>>>	RPR_02250
70	RPR_02280	>>>	45	382993	383113	>>>	120	336	>>>	RPR_02285
71	RPR_02300	>>>	1557	387055	387124	>>>	69	273	<<<	RPR_02325
72	RPR_02350	>>>	116	394856	394917	<<<	61	62	<<<	RPR_02355
73	RPR_02430	>>>	42	409691	409774	>>>	83	65	>>>	RPR_02435
74	RPR_02535	>>>	1700	428485	428616	>>>	131	-10	<<<	RPR_02555
75	RPR_02575	>>>	340	435785	436053	<<<	268	90	<<<	RPR_02585
76	RPR_02605	>>>	5	440581	440740	>>>	159	152	<<<	RPR_02610

77	RPR_02645	>>>	76	447274	447801	<<<	527	301	<<<	RPR_02655
78	RPR_02665	>>>	69	456181	456246	>>>	65	61	>>>	RPR_02670
79	RPR_02670	>>>	155	457393	457440	<<<	47	1311	<<<	RPR_02690
80	RPR_02670	>>>	865	458103	458506	>>>	403	245	<<<	RPR_02690
81	RPR_02825	>>>	1	477402	477524	>>>	122	-7	>>>	RPR_02830
82	RPR_02895	>>>	777	498267	498760	>>>	493	984	<<<	RPR_t07794
83	RPR_02930	>>>	1	501105	501201	>>>	96	156	>>>	RPR_02935
84	RPR_02955	<<<	1481	508482	508625	<<<	143	728	<<<	RPR_02990
85	RPR_03010	>>>	268	515262	515327	>>>	65	153	>>>	RPR_03020
86	RPR_03020	>>>	32	515981	516024	>>>	43	233	>>>	RPR_t07798
87	RPR_03035	>>>	131	519698	519826	>>>	128	93	>>>	RPR_03040
88	RPR_03050	>>>	-21	522190	522475	<<<	285	0	<<<	RPR_03060
89	RPR_03050	>>>	1	522212	522441	>>>	229	34	<<<	RPR_03060
90	RPR_03065	<<<	93	525495	525560	<<<	65	122	<<<	RPR_03070
91	RPR_03065	<<<	127	525529	525560	<<<	31	122	<<<	RPR_03070
92	RPR_03155	>>>	829	543637	543976	>>>	339	63	<<<	RPR_03170
93	RPR_03285	>>>	173	559244	559685	<<<	441	952	<<<	RPR_03305
94	RPR_03285	>>>	1386	560457	560600	>>>	143	37	<<<	RPR_03305
95	RPR_03285	>>>	1386	560457	560694	>>>	237	-57	<<<	RPR_03305
96	RPR_03530	>>>	1124	594191	594349	>>>	158	1793	>>>	RPR_t07804
97	RPR_t07804	>>>	2	596220	596314	>>>	94	17	<<<	RPR_03550
98	RPR_03560	>>>	19	598939	599039	>>>	100	124	>>>	RPR_03565
99	RPR_03575	>>>	18	600937	601078	>>>	141	50	>>>	RPR_03580
100	RPR_03585	>>>	41	603571	603632	>>>	61	1792	<<<	RPR_03605
101	RPR_03585	>>>	38	603568	603887	<<<	319	1537	<<<	RPR_03605
102	RPR_03760	>>>	359	630021	630131	<<<	110	1375	<<<	RPR_03775

103	RPR_03760	>>>	353	630015	630075	>>>	60	1431	<<<	RPR_03775
104	RPR_03760	>>>	1636	631298	631416	>>>	118	90	<<<	RPR_03775
105	RPR_03920	>>>	31	651925	651988	>>>	63	103	<<<	RPR_03925
106	RPR_03945	>>>	22	658391	658463	>>>	72	308	<<<	RPR_03950
107	RPR_03945	>>>	26	658395	658523	<<<	128	248	<<<	RPR_03950
108	RPR_03955	<<<	250	667520	667625	<<<	105	450	<<<	RPR_03960
109	RPR_03975	<<<	112	670549	670706	<<<	157	0	<<<	RPR_03980
110	RPR_04000	<<<	38	674971	675092	<<<	121	0	<<<	RPR_04005
111	RPR_04030	<<<	57	679687	679801	<<<	114	0	<<<	RPR_04035
112	RPR_04130	>>>	24	701736	701879	>>>	143	40	>>>	RPR_04135
113	RPR_04195	<<<	2263	714942	715042	<<<	100	1	<<<	RPR_04210
114	RPR_04260	>>>	1	724113	724200	>>>	87	94	>>>	RPR_04265
115	RPR_04280	>>>	944	731033	731066	>>>	33	1425	<<<	RPR_04315
116	RPR_04280	>>>	1700	731789	731867	>>>	78	624	<<<	RPR_04315
117	RPR_04390	>>>	153	751074	751179	>>>	105	4603	>>>	RPR_04435
118	RPR_04390	>>>	153	751074	751158	>>>	84	4624	>>>	RPR_04435
119	RPR_04390	>>>	4663	755584	755668	>>>	84	114	>>>	RPR_04435
120	RPR_04390	>>>	4663	755584	755638	>>>	54	144	>>>	RPR_04435
121	RPR_04475	<<<	84	765939	766025	<<<	86	18	<<<	RPR_04480
122	RPR_04590	>>>	4	782177	782282	>>>	105	81	<<<	RPR_04595
123	RPR_04660	<<<	98	796582	796631	<<<	49	704	<<<	RPR_04670
124	RPR_04690	<<<	1901	802990	803075	<<<	85	0	<<<	RPR_04710
125	RPR_04730	>>>	1818	808377	808865	>>>	488	5774	<<<	RPR_04800
126	RPR_04885	>>>	2	842288	842583	>>>	295	186	<<<	RPR_04890
127	RPR_04970	<<<	1116	862283	862424	<<<	141	32	<<<	RPR_04980
128	RPR_05020	>>>	293	872443	872814	<<<	371	1551	<<<	RPR_05055

130	RPR_05095	>>>	1	880897	880981	>>>	84	407	<<<	RPR_05100
131	RPR_05095	>>>	33	880929	881374	<<<	445	14	<<<	RPR_05100
132	RPR_05095	>>>	308	881204	881307	<<<	103	81	<<<	RPR_05100
133	RPR_05095	>>>	380	881276	881307	<<<	31	81	<<<	RPR_05100
134	RPR_05135	>>>	5121	892951	893022	>>>	71	112	>>>	RPR_05175
135	RPR_t07816	<<<	7433	908163	908202	<<<	39	94	<<<	RPR_05290
136	RPR_05320	<<<	485	915250	915792	<<<	542	2790	<<<	RPR_05355
137	RPR_05385	<<<	95	923009	923127	<<<	118	77	<<<	RPR_05390
138	RPR_t07818	<<<	2	937848	938022	>>>	174	72	<<<	RPR_05500
139	RPR_t07818	<<<	116	937962	938094	<<<	132	0	<<<	RPR_05500
140	RPR_05535	<<<	94	944039	944130	<<<	91	0	<<<	RPR_05540
141	RPR_05560	>>>	1	948840	949004	>>>	164	-21	>>>	RPR_05565
142	RPR_05575	>>>	1	950647	950716	>>>	69	79	>>>	RPR_05580
143	RPR_05580	>>>	149	951557	951631	>>>	74	180	>>>	RPR_05585
144	RPR_05600	<<<	108	955736	955789	<<<	53	197	<<<	RPR_05615
145	RPR_05735	<<<	65	974847	974917	<<<	70	42	<<<	RPR_05740
146	RPR_05735	<<<	101	974883	974917	<<<	34	42	<<<	RPR_05740
147	RPR_05790	<<<	102	988756	988868	<<<	112	0	<<<	RPR_05795
148	RPR_05805	<<<	100	991141	991247	<<<	106	0	<<<	RPR_05810
149	RPR_05890	<<<	-2	1005327	1005533	<<<	206	7	<<<	RPR_05895
150	RPR_06040	>>>	1	1026038	1026136	>>>	98	39	>>>	RPR_06045
151	RPR_06070	>>>	174	1033918	1034057	>>>	139	57	<<<	RPR_06080
152	RPR_06070	>>>	249	1033993	1034035	<<<	42	79	<<<	RPR_06080
153	RPR_06305	<<<	421	1064175	1064264	<<<	89	0	<<<	RPR_06310
154	RPR_06325	<<<	33	1066762	1066875	<<<	113	6	<<<	RPR_06330
155	RPR_06365	>>>	1493	1076427	1076552	>>>	125	2305	<<<	RPR_06410

156	RPR_06445	>>>	37	1084267	1084410	>>>	143	2501	<<<	RPR_06485
157	RPR_06445	>>>	140	1084370	1084487	<<<	117	2424	<<<	RPR_06485
158	RPR_06500	>>>	-4	1091693	1091846	<<<	153	0	<<<	RPR_06505
159	RPR_06605	>>>	38	1105868	1105989	>>>	121	198	<<<	RPR_06610
160	RPR_06730	>>>	1391	1132679	1132827	>>>	148	3278	<<<	RPR_06790
161	RPR_06730	>>>	2589	1133877	1134151	>>>	274	1954	<<<	RPR_06790
162	RPR_06795	>>>	104	1138568	1138654	<<<	86	4	<<<	RPR_06800
163	RPR_06945	>>>	2	1161873	1161947	>>>	74	112	>>>	RPR_06950
164	RPR_07030	<<<	102	1175936	1176056	<<<	120	34	<<<	RPR_07035
165	RPR_07085	>>>	76	1185416	1185494	<<<	78	1311	<<<	RPR_07100
166	RPR_07085	>>>	7	1185347	1185474	>>>	127	1331	<<<	RPR_07100
167	RPR_07180	>>>	1563	1203446	1203543	>>>	97	683	<<<	RPR_07220
168	RPR_07180	>>>	1563	1203446	1203507	>>>	61	719	<<<	RPR_07220
169	RPR_07180	>>>	1580	1203463	1203652	<<<	189	574	<<<	RPR_07220
170	RPR_07180	>>>	1563	1203446	1203667	>>>	221	559	<<<	RPR_07220
171	RPR_07180	>>>	1676	1203559	1203652	<<<	93	574	<<<	RPR_07220
172	RPR_07235	>>>	1	1211854	1211932	>>>	78	164	<<<	RPR_07240
173	RPR_07235	>>>	1	1211854	1211951	>>>	97	145	<<<	RPR_07240
174	RPR_07250	>>>	12	1214111	1214218	>>>	107	178	>>>	RPR_07255
175	RPR_07250	>>>	12	1214111	1214371	>>>	260	25	>>>	RPR_07255
176	RPR_07260	>>>	282	1216754	1216890	<<<	136	1367	<<<	RPR_07270
177	RPR_07260	>>>	1181	1217653	1218036	<<<	383	221	<<<	RPR_07270
178	RPR_07275	<<<	1690	1221907	1222038	<<<	131	292	<<<	RPR_07305
179	RPR_07305	<<<	633	1223360	1223421	<<<	61	0	<<<	RPR_07315
180	RPR_07330	>>>	776	1227561	1227899	<<<	338	401	<<<	RPR_07360
181	RPR_07460	<<<	1401	1244665	1244741	<<<	76	1	<<<	RPR_07480

182	RPR_t07838	<<<	91	1276490	1276582	<<<	92	0	<<<	RPR_07690
183	RPR_07695	<<<	126	1277232	1277325	<<<	93	0	<<<	RPR_07700
184	RPR_07700	<<<	969	1280128	1280633	<<<	505	819	<<<	RPR_07730
185	RPR_07700	<<<	2085	1281244	1281279	<<<	35	173	<<<	RPR_07730
186	RPR_05895	<<<	35	1006857	1006899	>>>	42	86	<<<	RPR_05900
187	RPR_06730	>>>	212	1131500	1131550	>>>	50	4555	<<<	RPR_06790
188	RPR_00180	<<<	139	29012	29089	<<<	77	594	>>>	RPR_00200
189	RPR_02280	>>>	45	382993	383076	>>>	83	373	>>>	RPR_02285
190	RPR_02825	>>>	1	477402	477494	>>>	92	23	>>>	RPR_02830
191	RPR_t07796	>>>	-38	504030	504270	<<<	240	130	>>>	RPR_02945
192	RPR_04980	<<<	1	863703	863742	>>>	39	137	<<<	RPR_04985

Rickettsia peacockii Plasmid pRPR SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RPR_p11	>>>	470	14750	14809	<<<	59	1003	>>>	RPR_p12
2	RPR_p11	>>>	309	14589	14809	<<<	220	1003	>>>	RPR_p12

Rickettsia rickettsii strain Iowa SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RrIowa_0535	<<<	1	445800	445857	>>>	57	49	<<<	RrIowa_0536
2	RrIowa_0707	<<<	12	585447	585507	>>>	60	138	>>>	RrIowa_0708
3	RrIowa_0757	>>>	1	623870	623962	>>>	92	22	>>>	RrIowa_0758
4	RrIowa_0902	<<<	82	733449	733911	>>>	462	-13	<<<	RrIowa_0904
5	RrIowa_1212	<<<	1	954409	954781	>>>	372	128	<<<	RrIowa_1213
7	RrIowa_1502	>>>	187	1189803	1190018	<<<	215	78	<<<	RrIowa_1503
8	RrIowa_1310	<<<	55	1033410	1033452	<<<	42	257	<<<	RrIowa_1312
9	RrIowa_1064	>>>	110	845704	845824	<<<	120	13	>>>	RrIowa_1065
10	RrIowa_0978	>>>	144	787122	787646	<<<	524	629	>>>	RrIowa_0979
11	RrIowa_0808	<<<	581	667860	667951	<<<	91	244	>>>	RrIowa_0809
12	RrIowa_0700	>>>	514	580211	580569	<<<	358	3	<<<	RrIowa_0701
13	RrIowa_0177	>>>	84	134491	134607	<<<	116	14	>>>	RrIowa_0178
14	RrIowa_0079	>>>	116	55252	55313	<<<	61	52	<<<	RrIowa_0080
15	RrIowa_0097	>>>	42	69858	69941	>>>	83	65	>>>	RrIowa_0098
16	RrIowa_0115	>>>	1	86371	86575	>>>	204	117	>>>	RrIowa_0116
17	RrIowa_0127	<<<	36	95132	95408	<<<	276	78	<<<	RrIowa_0128
18	RrIowa_0132	>>>	1	99920	100083	>>>	163	151	<<<	RrIowa_0133
19	RrIowa_0145	>>>	70	115879	115947	>>>	68	59	>>>	RrIowa_0146
20	RrIowa_0178	>>>	1	137041	137163	>>>	122	-7	>>>	RrIowa_0179
21	RrIowa_0222	>>>	377	181042	181395	>>>	353	215	>>>	RrIowa_0223
22	RrIowa_0224	>>>	1	190110	190509	>>>	399	-1	<<<	RrIowa_0225
23	RrIowa_0274	>>>	153	233349	233404	<<<	55	205	<<<	RrIowa_0275
24	RrIowa_0296	>>>	24	255696	256014	<<<	318	0	<<<	RrIowa_0297
25	RrIowa_0296	>>>	1	255673	255758	>>>	85	256	<<<	RrIowa_0297

26	RrIowa_0296	>>>	1	255673	255881	>>>	208	133	<<<	RrIowa_0297
27	RrIowa_0296	>>>	151	255823	256014	<<<	191	0	<<<	RrIowa_0297
28	RrIowa_0298	>>>	118	257667	257715	>>>	48	272	<<<	RrIowa_0299
29	RrIowa_0298	>>>	108	257657	257807	<<<	150	180	<<<	RrIowa_0299
30	RrIowa_0314	>>>	314	266973	267062	>>>	89	85	<<<	RrIowa_0315
31	RrIowa_0329	>>>	122	281303	281373	<<<	70	319	<<<	RrIowa_0330
32	RrIowa_0342	<<<	129	291617	291701	<<<	84	0	<<<	RrIowa_0343
33	RrIowa_0366	<<<	89	308630	308709	<<<	79	48	<<<	RrIowa_0367
34	RrIowa_0382	>>>	1	323748	323936	>>>	188	237	>>>	RrIowa_0383
35	RrIowa_0402	>>>	8	338204	338315	>>>	111	123	>>>	RrIowa_0403
36	RrIowa_0402	>>>	8	338204	338357	>>>	153	81	>>>	RrIowa_0403
37	RrIowa_0405	>>>	25	341593	341708	<<<	115	0	<<<	RrIowa_0406
38	RrIowa_0411	>>>	10	348779	348909	>>>	130	97	>>>	RrIowa_0412
39	RrIowa_0424	>>>	23	358874	359002	>>>	128	24	>>>	RrIowa_0425
40	RrIowa_0456	<<<	83	383022	383122	<<<	100	1	<<<	RrIowa_0457
41	RrIowa_0492	>>>	1	416439	416590	>>>	151	0	>>>	RrIowa_0493
42	RrIowa_0493	>>>	58	417894	418073	>>>	179	123	>>>	RrIowa_0494
43	RrIowa_0514	<<<	81	432793	432887	<<<	94	18	<<<	RrIowa_0515
44	RrIowa_0544	>>>	36	453754	453895	>>>	141	88	>>>	RrIowa_0545
45	RrIowa_0564	<<<	231	474287	474329	<<<	42	46	<<<	RrIowa_0565
46	RrIowa_0569	>>>	77	478993	479071	>>>	78	101	<<<	RrIowa_0572
47	RrIowa_0587	>>>	32	491042	491135	>>>	93	70	<<<	RrIowa_0588
48	RrIowa_0587	>>>	32	491042	491117	>>>	75	88	<<<	RrIowa_0588
49	RrIowa_0591	>>>	115	495186	495280	>>>	94	547	>>>	RrIowa_0592
50	RrIowa_0608	>>>	6	509056	509140	>>>	84	107	<<<	RrIowa_0609
51	RrIowa_0608	>>>	68	509118	509197	>>>	79	50	<<<	RrIowa_0609

52	RrIowa_0608	>>>	83	509133	509237	<<<	104	10	<<<	RrIowa_0609
53	RrIowa_0621	>>>	107	514601	514664	>>>	63	295	<<<	RrIowa_0622
54	RrIowa_0631	>>>	9	523803	524033	>>>	230	-25	>>>	RrIowa_0632
55	RrIowa_0653	>>>	8	540479	540609	>>>	130	77	>>>	RrIowa_0654
56	RrIowa_0663	>>>	57	554416	554498	>>>	82	40	>>>	RrIowa_0664
57	RrIowa_0673	>>>	6	561441	561537	>>>	96	104	<<<	RrIowa_0674
58	RrIowa_0673	>>>	38	561473	561568	<<<	95	73	<<<	RrIowa_0674
59	RrIowa_0673	>>>	6	561441	561517	>>>	76	124	<<<	RrIowa_0674
60	RrIowa_0705	<<<	29	583224	583339	<<<	115	0	<<<	RrIowa_0706
61	RrIowa_0710	>>>	1	588045	588173	>>>	128	1	>>>	RrIowa_0711
62	RrIowa_0718	<<<	67	592763	592849	<<<	86	0	<<<	RrIowa_0719
63	RrIowa_0721	<<<	31	596471	596627	<<<	156	0	<<<	RrIowa_0722
64	RrIowa_0731	<<<	95	601022	601105	<<<	83	0	<<<	RrIowa_0732
65	RrIowa_0733	>>>	25	603668	603828	>>>	160	61	>>>	RrIowa_0734
66	RrIowa_0736	>>>	34	606997	607130	>>>	133	33	>>>	RrIowa_0737
67	RrIowa_0748	>>>	1	616410	616547	>>>	137	37	<<<	RrIowa_0749
68	RrIowa_0757	>>>	1	623870	623985	>>>	115	-1	>>>	RrIowa_0758
69	RrIowa_0770	>>>	1	631056	631170	>>>	114	8	>>>	RrIowa_0773
70	RrIowa_0786	<<<	45	644659	644783	<<<	124	0	<<<	RrIowa_0787
71	RrIowa_0807	<<<	100	664468	664600	<<<	132	29	<<<	RrIowa_0808
72	RrIowa_0811	>>>	86	670433	670466	>>>	33	96	>>>	RrIowa_0812
73	RrIowa_0827	>>>	2	680486	680612	>>>	126	31	>>>	RrIowa_0828
74	RrIowa_0857	>>>	23	701691	701750	>>>	59	88	>>>	RrIowa_0858
75	RrIowa_0939	<<<	144	753947	754149	>>>	202	205	<<<	RrIowa_0941
76	RrIowa_0952	>>>	1	765543	765662	>>>	119	38	<<<	RrIowa_0953
77	RrIowa_0952	>>>	1	765543	765702	>>>	159	-2	<<<	RrIowa_0953

78	RrIowa_0962	>>>	1	773548	773616	>>>	68	51	>>>	RrIowa_0963
79	RrIowa_0974	<<<	54	784769	784872	<<<	103	0	<<<	RrIowa_0975
80	RrIowa_0978	>>>	44	787022	787178	>>>	156	1097	>>>	RrIowa_0979
81	RrIowa_1013	<<<	280	809304	809384	<<<	80	317	<<<	RrIowa_1014
82	RrIowa_1040	<<<	227	826126	826407	<<<	281	0	<<<	RrIowa_1041
83	RrIowa_1082	<<<	65	860966	861088	<<<	122	19	<<<	RrIowa_1084
84	RrIowa_1112	>>>	1	880772	881120	>>>	348	15	<<<	RrIowa_1113
85	RrIowa_1113	<<<	216	882615	882785	<<<	170	6	<<<	RrIowa_1114
86	RrIowa_1132	>>>	459	899263	899362	>>>	99	150	>>>	RrIowa_1133
87	RrIowa_1132	>>>	459	899263	899324	>>>	61	188	>>>	RrIowa_1133
88	RrIowa_1132	>>>	459	899263	899482	>>>	219	30	>>>	RrIowa_1133
89	RrIowa_1138	>>>	13	907667	907733	>>>	66	384	<<<	RrIowa_1139
90	RrIowa_1140	>>>	3	909890	910006	>>>	116	178	>>>	RrIowa_1141
91	RrIowa_1140	>>>	3	909890	910159	>>>	269	25	>>>	RrIowa_1141
92	RrIowa_1143	<<<	697	913465	913822	<<<	357	220	<<<	RrIowa_1144
93	RrIowa_1165	>>>	683	925840	926015	<<<	175	265	>>>	RrIowa_1166
94	RrIowa_1169	<<<	290	927558	927696	<<<	138	553	<<<	RrIowa_1170
95	RrIowa_1169	<<<	354	927622	927696	<<<	74	553	<<<	RrIowa_1170
96	RrIowa_1210	<<<	91	951490	951700	<<<	210	45	<<<	RrIowa_1211
97	RrIowa_1223	>>>	34	963846	963930	>>>	84	51	<<<	RrIowa_1224
98	RrIowa_1241	>>>	24	973982	974138	>>>	156	392	<<<	RrIowa_1242
99	RrIowa_1241	>>>	140	974098	974242	<<<	144	288	<<<	RrIowa_1242
100	RrIowa_1249	>>>	-4	981425	981578	<<<	153	0	<<<	RrIowa_1250
101	RrIowa_1271	>>>	174	995860	995982	>>>	122	200	<<<	RrIowa_1272
102	RrIowa_1271	>>>	342	996028	996174	<<<	146	8	<<<	RrIowa_1272
103	RrIowa_1293	<<<	423	1013993	1014061	<<<	68	0	<<<	RrIowa_1295

104	RrIowa_1301	>>>	1	1023632	1023790	>>>	158	-41	<<<	RrIowa_1302
105	RrIowa_1306	>>>	224	1027657	1027914	>>>	257	1	<<<	RrIowa_1307
106	RrIowa_1325	>>>	1	1040475	1040610	>>>	135	-11	>>>	RrIowa_1326
107	RrIowa_1334	>>>	1	1047380	1047730	>>>	350	64	>>>	RrIowa_1335
108	RrIowa_1346	>>>	57	1053781	1053906	<<<	125	127	<<<	RrIowa_1347
109	RrIowa_1366	>>>	1	1066062	1066182	>>>	120	38	<<<	RrIowa_1367
110	RrIowa_1394	>>>	123	1088453	1088540	>>>	87	70	>>>	RrIowa_1395
111	RrIowa_1400	<<<	118	1094224	1094346	<<<	122	110	<<<	RrIowa_1401
112	RrIowa_1440	<<<	57	1137022	1137085	<<<	63	0	<<<	RrIowa_1441
113	RrIowa_1449	<<<	67	1149405	1149491	<<<	86	0	<<<	RrIowa_1450
114	RrIowa_1468	>>>	117	1161335	1161417	>>>	82	1016	<<<	RrIowa_1470
115	RrIowa_1468	>>>	117	1161335	1161374	>>>	39	1059	<<<	RrIowa_1470
116	RrIowa_1502	>>>	45	1189661	1189740	>>>	79	356	<<<	RrIowa_1503
117	RrIowa_1502	>>>	415	1190031	1190096	<<<	65	0	<<<	RrIowa_1503
118	RrIowa_1528	>>>	42	1208377	1208485	>>>	108	754	>>>	RrIowa_1529
119	RrIowa_1589	<<<	85	1256835	1256927	<<<	92	0	<<<	RrIowa_1590
120	RrIowa_1595	>>>	1	1261194	1261410	>>>	216	-27	<<<	RrIowa_1596
121	RrIowa_1598	>>>	31	1264105	1264145	>>>	40	123	<<<	RrIowa_1599
122	RrIowa_0178	>>>	1	137041	137133	>>>	92	23	>>>	RrIowa_0179
123	RrIowa_0418	>>>	70	354154	354243	<<<	89	0	>>>	RrIowa_0419
124	RrIowa_0630	>>>	87	522436	522479	<<<	43	33	>>>	RrIowa_0631
125	RrIowa_0883	<<<	186	720213	720468	>>>	255	-38	<<<	RrIowa_0884
126	RrIowa_0969	<<<	16	782035	782141	<<<	106	50	<<<	RrIowa_0970

Rickettsia rickettsii strain Sheila Smith

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	A1G_00840	>>>	2615	153966	154095	>>>	129	92	>>>	A1G_00865
2	A1G_02535	<<<	1	445766	445823	>>>	57	49	<<<	A1G_02540
3	A1G_03335	<<<	12	585464	585524	>>>	60	138	>>>	A1G_03340
4	A1G_03580	>>>	1	623051	623143	>>>	92	22	>>>	A1G_03585
5	A1G_04290	<<<	1699	732781	733099	>>>	318	-13	<<<	A1G_04305
6	A1G_05625	<<<	13	943463	943835	>>>	372	128	<<<	A1G_05630
7	A1G_05980	>>>	152	998087	998270	>>>	183	-14	>>>	A1G_05985
9	A1G_07310	>>>	-24	1231351	1231582	<<<	231	0	<<<	A1G_07315
1	A1G_06100	>>>	1112	1021849	1022030	<<<	181	15	>>>	A1G_06105
11	A1G_04940	>>>	110	834328	834448	<<<	120	13	>>>	A1G_04945
12	A1G_03840	<<<	593	667011	667102	<<<	91	244	>>>	A1G_03845
13	A1G_00810	>>>	84	134392	134508	<<<	116	14	>>>	A1G_00815
14	A1G_00060	>>>	136	10332	10386	<<<	54	1110	<<<	A1G_00065
15	A1G_00275	>>>	55	44569	44777	<<<	208	126	<<<	A1G_00280
16	A1G_00380	>>>	116	55260	55321	<<<	61	52	<<<	A1G_00385
17	A1G_00470	>>>	42	69891	69974	>>>	83	65	>>>	A1G_00475
18	A1G_00560	>>>	1	86370	86574	>>>	204	601	>>>	A1G_00565
19	A1G_00600	>>>	47	95132	95408	<<<	276	78	<<<	A1G_00605
20	A1G_00625	>>>	1	99921	100084	>>>	163	151	<<<	A1G_00630
21	A1G_00680	>>>	70	115787	115855	>>>	68	59	>>>	A1G_00685
22	A1G_00815	>>>	1	136942	137064	>>>	122	-7	>>>	A1G_00820
23	A1G_00955	>>>	18	168895	168996	>>>	101	33	>>>	A1G_00960
24	A1G_00985	>>>	1	173565	173703	>>>	138	18	>>>	A1G_00990

25	A1G_01030	>>>	377	180949	181302	>>>	353	215	>>>	A1G_01035
26	A1G_01040	>>>	1	190017	190416	>>>	399	-1	<<<	A1G_01045
27	A1G_01260	>>>	153	233255	233310	<<<	55	205	<<<	A1G_01265
28	A1G_01380	>>>	24	255611	255929	<<<	318	0	<<<	A1G_01385
29	A1G_01380	>>>	1	255588	255673	>>>	85	256	<<<	A1G_01385
30	A1G_01380	>>>	151	255738	255929	<<<	191	0	<<<	A1G_01385
31	A1G_01380	>>>	1	255588	255796	>>>	208	133	<<<	A1G_01385
32	A1G_01390	>>>	108	257572	257673	<<<	101	0	<<<	A1G_01395
33	A1G_01390	>>>	1	257465	257630	>>>	165	43	<<<	A1G_01395
34	A1G_01475	>>>	314	266893	266982	>>>	89	86	<<<	A1G_01480
35	A1G_01615	<<<	129	291538	291622	<<<	84	0	<<<	A1G_01620
36	A1G_01700	>>>	1428	308557	308636	<<<	79	48	<<<	A1G_01725
37	A1G_t07539	>>>	1	323675	323863	>>>	188	237	>>>	A1G_01805
38	A1G_01905	>>>	8	338137	338248	>>>	111	123	>>>	A1G_01910
39	A1G_01905	>>>	8	338137	338290	>>>	153	81	>>>	A1G_01910
40	A1G_01920	>>>	25	341526	341641	<<<	115	0	<<<	A1G_01925
41	A1G_01950	>>>	10	348712	348842	>>>	130	142	>>>	A1G_01955
42	A1G_02015	>>>	23	358806	358934	>>>	128	24	>>>	A1G_02020
43	A1G_02165	<<<	83	382989	383089	<<<	100	1	<<<	A1G_02170
44	A1G_02350	>>>	58	417859	418038	>>>	179	165	>>>	A1G_02355
45	A1G_02445	<<<	84	432758	432852	<<<	94	18	<<<	A1G_02450
46	A1G_02575	>>>	36	453720	453861	>>>	141	88	>>>	A1G_02580
47	A1G_02675	<<<	258	474253	474295	<<<	42	46	<<<	A1G_02680
48	A1G_02695	>>>	487	478887	479037	>>>	150	101	<<<	A1G_02715
49	A1G_02795	>>>	32	491028	491121	>>>	93	70	<<<	A1G_02800
50	A1G_02795	>>>	32	491028	491103	>>>	75	88	<<<	A1G_02800

51	A1G_02870	>>>	1	503406	503504	>>>	98	39	>>>	A1G_02875
52	A1G_02900	>>>	6	509037	509121	>>>	84	107	<<<	A1G_02905
53	A1G_02900	>>>	83	509114	509218	<<<	104	10	<<<	A1G_02905
54	A1G_02900	>>>	68	509099	509178	>>>	79	50	<<<	A1G_02905
55	A1G_02960	>>>	274	514582	514645	>>>	63	295	<<<	A1G_02965
56	A1G_02995	>>>	9	523797	524024	>>>	227	-25	>>>	A1G_03000
57	A1G_03055	>>>	120	534170	534237	>>>	67	86	<<<	A1G_03060
58	A1G_03095	>>>	7	540472	540603	>>>	131	77	>>>	A1G_03100
59	A1G_03145	>>>	31	554385	554493	>>>	108	121	>>>	A1G_03150
60	A1G_03195	>>>	6	561436	561532	>>>	96	104	<<<	A1G_03200
61	A1G_03195	>>>	6	561436	561512	>>>	76	124	<<<	A1G_03200
62	A1G_03195	>>>	38	561468	561563	<<<	95	73	<<<	A1G_03200
63	A1G_03305	<<<	197	580584	580638	<<<	54	151	<<<	A1G_03310
64	A1G_03325	<<<	56	583241	583356	<<<	115	0	<<<	A1G_03330
65	A1G_03350	>>>	1	588062	588190	>>>	128	1	>>>	A1G_03355
66	A1G_03390	<<<	79	592779	592865	<<<	86	0	<<<	A1G_03395
67	A1G_03405	<<<	31	596487	596643	<<<	156	0	<<<	A1G_03410
68	A1G_03455	<<<	95	601038	601121	<<<	83	0	<<<	A1G_03460
69	A1G_03465	>>>	25	603684	603844	>>>	160	61	>>>	A1G_03470
70	A1G_03480	>>>	34	607013	607146	>>>	133	33	>>>	A1G_03485
71	A1G_03490	>>>	1	609833	609900	>>>	67	90	>>>	A1G_03495
72	A1G_03530	>>>	1	615589	615726	>>>	137	37	<<<	A1G_03535
73	A1G_03580	>>>	1	623051	623166	>>>	115	-1	>>>	A1G_03585
74	A1G_03635	>>>	1384	630233	630353	>>>	120	14	>>>	A1G_03655
75	A1G_03835	>>>	75	663619	663751	<<<	132	29	<<<	A1G_03840
76	A1G_03860	>>>	86	669584	669617	>>>	33	77	>>>	A1G_03865

77	A1G_03875	>>>	1466	673428	673695	>>>	267	9	>>>	A1G_03895
78	A1G_03940	>>>	2	679636	679762	>>>	126	31	>>>	A1G_03945
79	A1G_04040	>>>	63	693090	693173	>>>	83	82	<<<	A1G_04045
80	A1G_04090	>>>	23	700848	700907	>>>	59	166	>>>	A1G_04095
81	A1G_04110	<<<	95	707471	707520	<<<	49	876	<<<	A1G_04115
82	A1G_04280	<<<	4	730250	730500	<<<	250	0	<<<	A1G_04285
83	A1G_04450	>>>	564	753151	753353	>>>	202	219	<<<	A1G_04465
84	A1G_04525	>>>	1	764746	764865	>>>	119	38	<<<	A1G_04530
85	A1G_04525	>>>	1	764746	764905	>>>	159	-2	<<<	A1G_04530
86	A1G_04570	>>>	1	772760	772828	>>>	68	51	>>>	A1G_04575
87	A1G_04625	<<<	54	783987	784090	<<<	103	0	<<<	A1G_04630
88	A1G_04645	>>>	44	786240	786396	>>>	156	-35	<<<	A1G_04650
89	A1G_04720	>>>	111	796585	796694	<<<	109	0	<<<	A1G_04725
90	A1G_04725	<<<	280	797935	798015	<<<	80	317	<<<	A1G_04730
91	A1G_04745	<<<	44	802970	803268	<<<	298	0	<<<	A1G_04750
92	A1G_04840	<<<	227	814750	815031	<<<	281	0	<<<	A1G_t07567
93	A1G_t07571	>>>	1	831509	831604	>>>	95	17	<<<	A1G_04925
94	A1G_05030	<<<	65	849602	849724	<<<	122	19	<<<	A1G_05035
95	A1G_05160	>>>	1	869413	869761	>>>	348	15	<<<	A1G_05165
96	A1G_05165	<<<	216	871844	872014	<<<	170	6	<<<	A1G_05170
97	A1G_05250	>>>	459	888479	888578	>>>	99	150	>>>	A1G_05255
98	A1G_05250	>>>	459	888479	888540	>>>	61	188	>>>	A1G_05255
99	A1G_05250	>>>	459	888479	888698	>>>	219	30	>>>	A1G_05255
10	A1G_05280	>>>	13	896883	896949	>>>	66	164	<<<	A1G_05285
101	A1G_05300	>>>	3	899107	899223	>>>	116	178	>>>	A1G_05305
102	A1G_05300	>>>	3	899107	899376	>>>	269	25	>>>	A1G_05305

103	A1G_05315	<<<	699	902684	903041	<<<	357	220	<<<	A1G_05320
104	A1G_05385	>>>	131	912537	912661	>>>	124	432	<<<	A1G_05400
105	A1G_05420	<<<	402	916778	916916	<<<	138	553	<<<	A1G_t07573
106	A1G_05420	<<<	466	916842	916916	<<<	74	553	<<<	A1G_t07573
107	A1G_05615	<<<	91	940545	940755	<<<	210	45	<<<	A1G_05620
108	A1G_05635	<<<	33	946513	946626	<<<	113	5	<<<	A1G_05640
109	A1G_05765	>>>	24	963029	963185	>>>	156	264	>>>	A1G_05775
110	A1G_05820	>>>	-4	970472	970625	<<<	153	0	<<<	A1G_05825
111	A1G_05875	>>>	2834	982324	982550	>>>	226	190	>>>	A1G_05905
112	A1G_05920	>>>	174	984298	984420	>>>	122	200	<<<	A1G_05925
113	A1G_05920	>>>	342	984466	984612	<<<	146	8	<<<	A1G_05925
114	A1G_06015	<<<	1246	1002431	1002499	<<<	68	0	<<<	A1G_06030
115	A1G_06060	>>>	1	1012070	1012228	>>>	158	-41	<<<	A1G_06065
116	A1G_06165	>>>	1	1028913	1029048	>>>	135	-11	>>>	A1G_06170
117	A1G_t07583	>>>	1	1035818	1036168	>>>	350	314	<<<	A1G_06210
118	A1G_06265	>>>	57	1042220	1042345	<<<	125	127	<<<	A1G_06270
119	A1G_06375	>>>	1	1054495	1054615	>>>	120	38	<<<	A1G_06380
120	A1G_06510	>>>	123	1076882	1076969	>>>	87	70	>>>	A1G_06515
121	A1G_06540	<<<	118	1082653	1082692	<<<	39	193	<<<	A1G_06545
122	A1G_06735	<<<	57	1125452	1125515	<<<	63	0	<<<	A1G_06740
123	A1G_06780	<<<	67	1137835	1137921	<<<	86	0	<<<	A1G_06785
124	A1G_06880	>>>	117	1149921	1150003	>>>	82	1060	>>>	A1G_06890
125	A1G_06880	>>>	117	1149921	1149960	>>>	39	1103	>>>	A1G_06890
126	A1G_07040	>>>	45	1179131	1179210	>>>	79	356	<<<	A1G_07045
127	A1G_07040	>>>	415	1179501	1179566	<<<	65	0	<<<	A1G_07045
128	A1G_t07593	<<<	85	1246358	1246450	<<<	92	0	<<<	A1G_07460

129	A1G_07480	>>>	1	1250716	1250932	>>>	216	-27	<<<	A1G_07485
130	A1G_07495	>>>	3	1253599	1253667	>>>	68	1161	<<<	A1G_07510
131	A1G_00815	>>>	1	136942	137034	>>>	92	23	>>>	A1G_00820
132	A1G_01215	>>>	222	228632	228671	>>>	39	437	<<<	A1G_01220
133	A1G_02050	>>>	45	364956	365040	>>>	84	179	>>>	A1G_02055
134	A1G_02990	>>>	87	522430	522473	<<<	43	33	>>>	A1G_02995
135	A1G_04195	<<<	582	719400	719655	>>>	255	-38	<<<	A1G_t07559
136	A1G_04605	<<<	16	781253	781359	<<<	106	50	<<<	A1G_04610

Rickettsia prowazekii strain Madrid E SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RP194	>>>	1599	234682	235015	>>>	333	4231	<<<	RP196
2	RP490	<<<	333	605623	605708	>>>	85	381	>>>	RP492
3	RPt29a	<<<	289	963322	963406	>>>	84	17	<<<	RP771
4	RP779	<<<	112	976666	976735	<<<	69	63	>>>	RP780
5	RP098	>>>	413	110032	110133	>>>	101	1167	>>>	RP099
6	RP106	>>>	1	124861	124991	>>>	130	29	>>>	RP107
7	RP139	>>>	205	160002	160087	>>>	85	562	>>>	RP140
8	RP194	>>>	5916	238999	239036	<<<	37	210	<<<	RP196
9	RP254	<<<	48	312195	312320	<<<	125	51	<<<	RP255
10	RP324	>>>	1	397342	397695	>>>	353	861	<<<	RP325
11	RP375	>>>	922	460896	461178	>>>	282	340	<<<	RP376
12	RP376	<<<	686	463150	463296	<<<	146	2	<<<	RP377
13	RP384	>>>	685	472590	472726	>>>	136	181	>>>	RP385
14	RP413	>>>	1045	506319	506409	>>>	90	1931	<<<	RP414
15	RP419	>>>	2	516831	517043	>>>	212	150	>>>	RP420
16	RP434	<<<	61	532298	532509	<<<	211	2	<<<	RP435
17	RP442	>>>	64	540952	541129	>>>	177	411	>>>	RP443
18	RP517	<<<	18	640871	641007	<<<	136	210	>>>	RPt16
19	RP554	<<<	18	686834	686985	<<<	151	1143	<<<	RP555
20	RP560	<<<	1305	702193	702392	<<<	199	93	<<<	RP561
21	RP573	<<<	1418	721583	721733	<<<	150	399	<<<	RPt20
22	RP601	>>>	536	758794	758998	>>>	204	896	>>>	RP602
23	RPr03	<<<	366	774136	774195	<<<	59	1023	<<<	RP614
24	RP675	>>>	80	839681	839875	>>>	194	802	>>>	RP676

25	RP703	>>>	1648	881002	881242	>>>	240	21	<<<	RP704
27	RP723	<<<	97	912290	912349	<<<	59	214	<<<	RP724

Rickettsia prowazekii strain Breinl SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	H375_4260	>>>	434	534507	534174	<<<	333	1,932	<<<	H375_4280
2	H375_1190	<<<	466	163641	163556		85	418	>>>	H375_1200
3	H375_7520	>>>	101	915603	915519	<<<	84	373	>>>	H375_t110
4	H375_7430	<<<	63	902196	902265	>>>	69	112	>>>	H375_7440
5	H375_5310	<<<	1,268	659158	659057	<<<	101	514	<<<	H375_5320
6	H375_5220	<<<	204	644329	644199	<<<	130	131	<<<<	H375_5230
7	H375_4260	>>>	434	534507	534174	<<<	333	1,932	<<<	H375_4280
8	H375_4230	>>>	210	530135	530172	>>>	37	1,514	<<<	H375_4240
9	H375_3620	>>>	51	456876	457001	>>>	125	48	>>>	H375_3630
10	H375_2870	<<<	499	371859	371506	<<<	353	354	<<<	H375_2880
11	H375_2330	>>>	622	308324	308042	<<<	282	391	>>>	H375_2340
12	H375_2320	>>>	2	305924	306070	>>>	146	686	>>>	H375_2330
13	H375_2260	<<<	96	296630	296494	<<<	136	821	<<<	H375_2270
14	H375_1950	>>>	1,966	262900	262810	<<<	90	1,135	<<<	H375_1960
15	H375_1890	<<<	362	252443	252231	<<<	212	214	<<<	H375_1900
16	H375_1740	>>>	2	236766	236976	>>>	210	61	>>>	H375_1750
17	H375_1660	<<<	588	228312	228135	<<<	177	241	<<<	H375_1670
18	H375_t290	<<<	210	128257	128393	>>>	136	18	>>>	H375_950
19	H375_590	>>>	1,143	82284	82435	>>>	151	18	>>>	H375_600
20	H375_540	>>>	-149	66876	67075	>>>	199	49	>>>	H375_550
21	H375_t10	>>>	395	47542	47692	>>>	150	1,418	>>>	H375_410
22	H375_120	<<<	1,100	10482	10278	<<<	204	740	<<<	H375_130
23	H375_9210	>>>	1,082	1105018	1104959	<<<	59	425	>>>	H375_r20
24	H375_8560	<<<	512	1039473	1039278	<<<	195	275	<<<	H375_8570
25	H375_8270	>>>	261	998167	997927	<<<	240	1,888	<<<	H375_8280

27	H375_7980	>>>	273	966633	966574	<<<	59	156	>>>	H375_7990

Rickettsia typhi SIPHT Predictions

sRNAName	UpGENEnumber	Direction	Distance	sRNAstart	sRNAend	Direction	Length	Distance	DnGENEdir.	DnGENEnumber
1	RT0298	>>>	448	379632	379747	>>>	115	1201	>>>	RT0299
2	RT0476	<<<	336	608966	609053	>>>	87	384	>>>	RT0478
3	RT0663	>>>	204	839741	840183	>>>	442	1353	>>>	RT0664
4	RT0757	<<<	285	961257	961342	>>>	85	19	<<<	RT0758
5	RT0563	<<<	229	729224	729549	<<<	325	326	<<<	RT0564
6	RT0038	<<<	920	50655	50954	<<<	299	62	<<<	RT0039
7	RT0118	>>>	1	153571	153713	>>>	142	16	>>>	RT0119
8	RT0213	<<<	1026	276200	276466	<<<	266	68	<<<	RT0214
9	RT0246	<<<	42	310544	310668	<<<	124	6	<<<	RT0247
10	RT0421	<<<	58	532921	533148	<<<	227	8	<<<	RT0422
11	RT0429	>>>	464	542199	542289	>>>	90	75	>>>	RT0430
12	RT0457	>>>	67	587356	587475	>>>	119	205	>>>	RT0458
13	RT0559	>>>	1	724753	724945	>>>	192	48	<<<	RT0560
14	RT0710	<<<	96	908099	908160	<<<	61	243	<<<	RT0711
15	RT0719	>>>	161	916472	916839	>>>	367	653	>>>	RT0720

Additional File 2. Common Target Genes.

sRNA candidate name	Total number of target genes predicted by TargetRNA2	Total number of target genes predicted by CopraRNA	Number of common targets predicted by both programs	List of target genes predicted by both programs	Protein name
1	39	47	5	H375_7250	NADH dehydrogenase subunit G
				H375_7510	3-oxoacyl-[acyl-carrier-protein] synthase 3
				H375_8730	50S ribosomal protein L3
				H375_8330	Bicyclomycin resistance protein
				H375_5250	hypothetical protein
2	3	43	2	H375_8960	30S ribosomal protein S13
				H375_1610	hypothetical protein
3	4	40	2	H375_1880	3'-5' exonuclease
				H375_1280	hypothetical protein
4	11	23	0		
5	27	50	2	H375_1570	Isopentenyl-diphosphate delta-isomerase
				H375_2460	Apolipoprotein N-acyltransferase
6	21	49	6	H375_480	Cell division protein FtsL
				H375_1770	Succinyl-CoA ligase [ADP-forming] subunit alpha-2
				H375_6340	Coproporphyrinogen-III oxidase aerobic
				H375_8550	Exodeoxyribonuclease III
				H375_390	Protein translocase subunit SecA
				H375_3220	Hypothetical protein
7	0	53	0		
8	8	37	0		
9	19	43	0		
10	32	49	5	H375_5120	GTPase Era
				H375_9140	2-acylglycerophosphoethanolamine acyltransferase
				H375_4430	2-oxoglutarate dehydrogenase E1 component
				H375_1600	Lon protease
				H375_230	hypothetical protein
11	17	49	2	H375_4380	protease DO
				H375_3930	hypothetical protein
12	11	44	1	H375_3680	Cell division protein FtsQ
13	17	49	3	H375_6140	190 kDa antigen

				H375_3520	Isocitrate dehydrogenase
				H375_6960	hypothetical protein
14	13	37	2	H375_6840	Single-stranded DNA-binding protein
				H375_5400	hypothetical protein
15	25	50	5	H375_1110	tRNA pseudouridine synthase B
				H375_3700	UDP-N-acetylenolpyruvoylglucosamine reductase
				H375_3020	Guanosine polyphosphate pyrophosphohydrolase/synthetase
				H375_9070	Guanosine-3',5'-bis(Diphosphate) 3'-pyrophosphohydrolase
				H375_2300	hypothetical protein
16	12	45	2	H375_7470	DNA polymerase I
				H375_2240	Holliday junction DNA helicase RuvB
17	18	51	3	H375_4200	Chaperone protein HscA
				H375_8970	DNA-directed RNA polymerase subunit alpha
				H375_8730	50S ribosomal protein L3
18	2	54	0		
19	9	45	1	H375_3150	sugar phosphate isomerase
20	10	39	0		
21	12	39	0		
22	14	46	6	H375_1000	Ribonucleoside-diphosphate reductase
				H375_1850	Phosphatidate cytidylyltransferase
				H375_4260	Bifunctional penicillin-binding protein 1C
				H375_6700	Sua5
				H375_3980	Threonine--tRNA ligase
				H375_7980	hypothetical protein
23	10	45	0		
24	24	54	4	H375_7950	Dephospho-CoA kinase
				H375_2830	Thioredoxin peroxidase 1
				H375_3260	VirB10 protein
				H375_4700	hypothetical protein
25	29	36	0		
27	6	37	0		

Additional File 3. List of small RNAs found within the *Rickettsia prowazekii* genome.

Rp_sR	Approximate Start	Approximate Stop	Size (bp)	Strand	Type of sRNA	Homology	Strand Orientation	Notes
1	10482	10278	204	R	Trans	TG	</</<	Predicted as SIPHT #22
2	10774	11003	230	F	Trans	Rp	</>/<	Identified by RNA-seq
3	14215	13984	232	R	Cis	Rp	H375_160	Identified by RNA-seq
4	15459	15203	257	R	Cis	TG, TRG, SFG	H375_160	Identified by RNA-seq
5	19850	19622	229	R	Trans	TG, TRG	>/</>	Identified by RNA-seq
6	22107	22298	192	F	Trans	Rp	</>/>	Identified by RNA-seq
7	24935	25206	272	F	Cis	TG, TRG, SFG	H375_230	Identified by RNA-seq
8	47692	47542	150	F	Trans	Rp	>/>/>	Predicted as SIPHT #21
9	48035	48200	166	F	Trans	Rp	>/>/>	Identified by RNA-seq
10	48620	48834	215	F	Trans	Rp	>/>/>	Identified by RNA-seq
11	71739	72043	305	F	Cis	TG, TRG, SFG	H375_570	Identified by RNA-seq
12	76189	76452	264	F	Cis	AG, TG, TRG	H375_570	Identified by RNA-seq
13	77115	77335	221	F	Trans	Rp	</>/<	Identified by RNA-seq
14	78259	78513	255	F	Trans	Rp	</>/>	Identified by RNA-seq
15	88423	88616	194	F	Trans	Rp	>/>/<	Identified by RNA-seq
16	100282	99698	585	R	Cis	TG, TRG, SFG	H375_740	Identified by RNA-seq
17	105748	105486	263	R	Trans	Rp	>/</>	Identified by RNA-seq
18	116363	116200	164	R	Cis	TG	H375_870	Identified by RNA-seq
19	132593	132781	189	F	Trans	Rp	</>/>	Identified by RNA-seq
20	163641	163556	85	R	Trans	AG, TG	</</>	Predicted as SIPHT #2
21	173793	173974	182	F	Cis	TG, TRG, SFG	H375_1300	Identified by RNA-seq
22	177789	178118	330	F	Cis	TG, SFG	H375_1330	Identified by RNA-seq
23	185267	185484	218	F	Cis	TG, TRG, SFG	H375_1400	Identified by RNA-seq
24	218432	218657	226	F	Cis	TG, TRG, SFG	H375_1610	Identified by RNA-seq
25	257587	257856	270	F	Cis	TG, TRG, SFG	H375_1930	Identified by RNA-seq
26	261134	261341	208	F	Trans	Rp	>/>/<	Identified by RNA-seq
27	262322	262601	280	F	Trans	Rp	>/>/<	Identified by RNA-seq
28	275336	275043	294	R	Cis	TG, TRG	H375_2050	Identified by RNA-seq
29	306070	305924	146	F	Trans	Rp	>/>/>	Predicted as SIPHT #12
30	308324	308042	282	R	Trans	TG	>/</<	Predicted as SIPHT #11
31	308652	308848	197	F	Trans	Rp	>/>/<	Identified by RNA-seq
32	309308	309564	257	F	Cis	TG, TRG, SFG	H375_2350	Identified by RNA-seq
33	315408	315772	365	F	Cis	TG, TRG, SFG	H375_2380	Identified by RNA-seq
34	326339	326637	299	F	Cis	TG, TRG, SFG	H375_2470	Identified by RNA-seq

35	334531	334333	199	R	Cis	TG, TRG, SFG	H375_2560	Identified by RNA-seq
36	342505	342292	214	R	Cis	TG, TRG, SFG	H375_2630	Identified by RNA-seq
37	344054	343930	125	R	Trans	Rp	>/</<	Identified by RNA-seq
38	366166	366242	77	F	Trans	Rp	>/>/>	Identified by RNA-seq
39	371859	371506	353	R	Trans	TG	>/</<	Predicted as SIPHT #10
40	373780	373983	204	F	Cis	TG, TRG, SFG	H375_2910	Identified by RNA-seq
41	378222	378333	112	F	Cis	TG, TRG, SFG	H375_2940	Identified by RNA-seq
42	407851	407640	212	R	Trans	Rp	>/</<	Identified by RNA-seq
43	427133	426907	227	R	Trans	Rp	>/</>	Identified by RNA-seq
44	457001	456876	125	F	Trans	TG	>/>/>	Predicted as SIPHT #9
45	458370	458555	227	R	Trans	TG	>/</>	Identified by RNA-seq
46	462025	462293	268	F	Cis	TG, TRG, SFG	H375_3670	Identified by RNA-seq
47	481518	481833	316	F	Cis	TG, TRG, SFG	H375_3890	Identified by RNA-seq
48	494600	494387	214	R	Trans	Rp	</</<	Identified by RNA-seq
49	504250	504506	256	F	Trans	Rp	>/>/>	Identified by RNA-seq
50	514132	513893	240	R	Trans	TG	</</<	Identified by RNA-seq
51	531332	531187	146	R	Trans	Rp	>/</<	Identified by RNA-seq
52	535456	535256	201	R	Trans	Rp	>/</<	Identified by RNA-seq
53	558930	558719	212	R	Trans	Rp	>/</<	Identified by RNA-seq
54	644329	644199	130	R	Trans	Rp	</</<	Predicted as SIPHT #6
55	659164	659057	107	R	Trans	Rp	</</<	Predicted as SIPHT #5
56	669521	669772	252	F	Trans	Rp	>/>/<	Identified by RNA-seq
57	697106	696934	173	R	Trans	Rp	>/</<	Identified by RNA-seq
58	812662	812412	251	R	Trans	Rp	</</<	Identified by RNA-seq
59	830223	830416	194	F	Trans	Rp	>/>/<	Identified by RNA-seq
60	844283	844606	324	F	Trans	Rp	>/>/>	Identified by RNA-seq
61	859237	859436	200	F	Trans	Rp	>/>/>	Identified by RNA-seq
62	909871	910281	411	F	Cis	TG, TRG, SFG	H375_7470	Identified by RNA-seq
63	925625	925929	305	F	Trans	Rp	</>/<	Identified by RNA-seq
64	958095	957827	269	R	Trans	Rp	</</<	Identified by RNA-seq
65	968393	968664	272	F	Trans	Rp	>/>/>	Identified by RNA-seq
66	973818	973352	467	R	Trans	TG	>/</>	Identified by RNA-seq
67	998167	997927	240	R	Trans	Rp	>/</<	Predicted as SIPHT #25
68	1039473	1039278	195	R	Trans	Rp	</</<	Predicted as SIPHT #24
69	1046344	1046671	328	F	Trans	Rp	</>/<	Identified by RNA-seq
70	1105018	1104959	59	F	Trans	Rp	>/>/>	Predicted as SIPHT #23

[a]RNA sequencing data demonstrates that *R. prowazekii* strain Breinl encodes for at least 58 candidate sRNAs. Each of these candidates are listed here. The sRNA number, the approximate start location based on the start of the RNA sequencing reads, the approximate stop location based again on the RNA sequencing reads, the nucleotide size, the sRNA carrying strand, the nature of sRNA, and its to other *Rickettsia* species are shown. The column labeled as "Strand Orientation" refers to the orientation of the upstream gene, the trans-acting sRNA, and downstream gene. For cis-acting sRNAs, the corresponding ORF is listed in this column (F=forward; R=reverse; Rp=*R. prowazekii* only; TG=typhus group (both *R. prowazekii* and *R. typhi*); TRG=transitional group; SFG=spotted fever group; > = sense strand; < = anti-sense strand). Sequences may not be found in all species of a particular group.

Additional File 4. Complete list of predicted targets by TargetRNA2 and IntaRNA.

The complete list of predicted targets by TargetRNA2 and IntaRNA for *Rp*_sR17, *Rp*_sR60, and *Rp*_sR67. Only those targets with a P≤0.05 were included for analysis.

*Rp*_sR17

TargetRNA2

Rank	Synonym	Energy	Pvalue	sRNA_start	sRNA_stop	mRNA_start	mRNA_stop
1	H375_4030	-12.25	0.006	127	141	7	20
2	H375_4670	-11.51	0.01	86	100	-10	5
3	H375_1770	-10.95	0.014	76	92	-79	-62
4	H375_1820	-10.67	0.017	106	119	-60	-47
5	H375_8570	-10.25	0.021	89	100	-67	-56
6	H375_1570	-9.97	0.024	18	31	-65	-52
7	H375_1330	-9.89	0.025	138	146	-80	-72
8	H375_4570	-9.73	0.027	4	18	-63	-49
9	H375_6540	-9.43	0.031	115	128	-75	-62
10	H375_4140	-9.39	0.031	138	150	-73	-61
11	H375_9000	-9.23	0.034	34	48	-3	11
12	H375_5880	-8.99	0.037	91	99	-78	-70
13	H375_1940	-8.95	0.038	14	29	-77	-61
14	H375_7090	-8.67	0.043	36	44	4	12
15	H375_290	-8.64	0.043	35	48	-50	-38
16	H375_3060	-8.64	0.043	112	128	-49	-32
17	H375_5480	-8.53	0.045	111	121	-68	-58
18	H375_6750	-8.49	0.046	119	132	-27	-14
19	H375_8160	-8.38	0.048	161	170	2	11

IntaRNA

p-value	fdr value	Target	Position	Position	Energy
0.000292	0.26862	H375_7580	92 -- 148	43 -- 86	-15.4219
0.00182	0.40004	H375_5880	95 -- 114	37 -- 54	-13.4649
0.001938	0.40004	H375_4030	71 -- 102	121 -- 154	-13.3933
0.0020459	0.40004	H375_900	140 -- 148	39 -- 47	-13.3313
0.0021741	0.40004	H375_5240	7 -- 36	27 -- 72	-13.2615
0.0026978	0.41366	H375_5900	131 -- 144	4 -- 17	-13.0113
0.0035863	0.47135	H375_6990	121 -- 148	71 -- 100	-12.6755
0.0048452	0.5572	H375_3580	28 -- 40	5 -- 17	-12.3134
0.0092242	0.85907	H375_4740	136 -- 146	103 -- 113	-11.5124
0.0097439	0.85907	H375_4400	93 -- 135	55 -- 103	-11.4425
0.0113355	0.85907	H375_150	66 -- 77	102 -- 113	-11.2481
0.0113645	0.85907	H375_7180	91 -- 103	3 -- 15	-11.2448
0.012139	0.85907	H375_290	25 -- 34	6 -- 15	-11.1594
0.0149482	0.95753	H375_6880	6 -- 24	153 -- 173	-10.887
0.0208462	0.95753	H375_6680	60 -- 83	29 -- 48	-10.4428
0.0222944	0.95753	H375_1740	17 -- 56	13 -- 61	-10.3517
0.0224259	0.95753	H375_4750	99 -- 107	5 -- 13	-10.3437

0.0225713	0.95753	H375_8450	120 -- 130	5 -- 15	-10.3349
0.0241892	0.95753	H375_5420	5 -- 13	4 -- 12	-10.2404
0.0268507	0.95753	H375_4570	15 -- 26	5 -- 16	-10.0969
0.0278154	0.95753	H375_5050	114 -- 132	4 -- 21	-10.0481
0.0301504	0.95753	H375_5680	82 -- 92	5 -- 15	-9.93612
0.0307797	0.95753	H375_6370	131 -- 141	46 -- 56	-9.9073
0.031912	0.95753	H375_2400	18 -- 34	156 -- 171	-9.85678
0.0320834	0.95753	H375_8410	43 -- 61	27 -- 45	-9.84928
0.0336721	0.95753	H375_5790	53 -- 62	33 -- 42	-9.78142
0.0349641	0.95753	H375_8080	130 -- 146	1 -- 16	-9.72836
0.0368166	0.95753	H375_9020	106 -- 146	5 -- 42	-9.65532
0.0377172	0.95753	H375_3690	53 -- 73	27 -- 47	-9.62102
0.0383444	0.95753	H375_5440	136 -- 149	100 -- 113	-9.59757
0.0385662	0.95753	H375_8180	128 -- 136	5 -- 13	-9.58936
0.0397341	0.95753	H375_5060	108 -- 138	41 -- 75	-9.54683
0.0406908	0.95753	H375_8870	34 -- 102	4 -- 64	-9.51283
0.0409879	0.95753	H375_5540	128 -- 145	37 -- 54	-9.50242
0.0427328	0.95753	H375_2270	122 -- 139	150 -- 167	-9.44263
0.0428752	0.95753	H375_2490	110 -- 140	139 -- 167	-9.43785
0.0440463	0.95753	H375_2540	60 -- 85	25 -- 49	-9.39907
0.0441044	0.95753	H375_3590	12 -- 49	39 -- 75	-9.39717
0.045589	0.95753	H375_2940	143 -- 149	38 -- 44	-9.34939
0.0489416	0.95753	H375_3030	93 -- 134	26 -- 65	-9.24648

*Rp*_sR60

TargetRNA2

Rank	Synonym	Energy	Pvalue	sRNA_start	sRNA_stop	mRNA_start	mRNA_stop
1	H375_7590	-12.13	0.007	79	92	-35	-22
2	H375_4430	-10.49	0.018	80	93	-37	-24
3	H375_4980	-10.45	0.019	63	75	-20	-8
4	H375_1260	-10.27	0.02	22	37	-16	-1
5	H375_3340	-10.18	0.021	5	16	-49	-38
6	H375_1680	-10.14	0.022	7	16	10	19
7	H375_3020	-9.93	0.024	83	97	-44	-30
8	H375_7240	-9.88	0.025	11	22	-4	8
9	H375_2400	-9.84	0.025	36	49	-78	-65
10	H375_4120	-9.73	0.027	5	19	-1	15
11	H375_80	-9.49	0.03	32	45	-60	-47
12	H375_1850	-9.39	0.031	39	51	-2	11
13	H375_3390	-9.34	0.032	11	25	-7	8
14	H375_4320	-9.15	0.035	90	98	-47	-39
15	H375_7400	-9.07	0.036	37	45	-76	-68
16	H375_1880	-9.06	0.036	77	93	-6	15
17	H375_8280	-8.99	0.037	79	89	-16	-6
18	H375_1840	-8.92	0.039	6	17	-53	-43

19	H375_6310	-8.8	0.041	22	34	6	18
20	H375_8880	-8.73	0.042	59	74	-66	-52
21	H375_4660	-8.68	0.043	39	51	-44	-32
22	H375_4140	-8.58	0.044	71	89	-80	-63
23	H375_1220	-8.52	0.045	42	55	-69	-56

IntaRNA

p-value	fdr value	Target	Position	Position	Energy
0.0004006	0.23136	H375_7670	132 -- 150	79 -- 99	-16.4559
0.0005074	0.23136	H375_7550	66 -- 92	72 -- 99	-16.1203
0.0012762	0.38796	H375_8740	11 -- 41	68 -- 93	-14.7801
0.0031426	0.71652	H375_6080	39 -- 96	47 -- 99	-13.4238
0.0051069	0.74743	H375_4620	132 -- 146	83 -- 97	-12.6729
0.0051886	0.74743	H375_6370	117 -- 144	70 -- 97	-12.6481
0.0057368	0.74743	H375_5790	127 -- 149	54 -- 76	-12.4908
0.0067875	0.77377	H375_4980	57 -- 67	64 -- 74	-12.226
0.0095703	0.91081	H375_8880	11 -- 26	58 -- 73	-11.6792
0.0116703	0.91081	H375_5600	118 -- 150	74 -- 99	-11.3598
0.0121839	0.91081	H375_5090	126 -- 144	82 -- 97	-11.2901
0.0130033	0.91081	H375_1850	14 -- 99	27 -- 99	-11.1845
0.0131976	0.91081	H375_3420	21 -- 35	80 -- 94	-11.1604
0.014019	0.91081	H375_5930	35 -- 56	77 -- 99	-11.0621
0.0149804	0.91081	H375_8110	103 -- 115	80 -- 93	-10.9538
0.016865	0.94584	H375_300	1 -- 38	62 -- 97	-10.7595
0.0187961	0.94584	H375_4840	16 -- 28	80 -- 92	-10.5808
0.0232714	0.94584	H375_5870	117 -- 127	63 -- 74	-10.226
0.0234785	0.94584	H375_9210	1 -- 14	77 -- 90	-10.2112
0.0247598	0.94584	H375_7620	103 -- 125	80 -- 99	-10.1223
0.0254668	0.94584	H375_3880	124 -- 144	77 -- 96	-10.0751
0.0271228	0.94584	H375_3020	33 -- 45	84 -- 97	-9.96923
0.0298539	0.94584	H375_2040	128 -- 148	68 -- 84	-9.80733
0.0313134	0.94584	H375_7590	42 -- 53	80 -- 91	-9.72647
0.0316496	0.94584	H375_6720	118 -- 133	80 -- 100	-9.70835
0.0320754	0.94584	H375_9010	2 -- 15	86 -- 97	-9.68566
0.0331266	0.94584	H375_1340	118 -- 127	61 -- 70	-9.63084
0.0337198	0.94584	H375_650	83 -- 99	83 -- 97	-9.60063
0.0337451	0.94584	H375_50	139 -- 150	88 -- 99	-9.59935
0.0357489	0.94584	H375_5950	92 -- 104	86 -- 98	-9.50094
0.0359822	0.94584	H375_4070	131 -- 148	70 -- 92	-9.48982
0.0360294	0.94584	H375_7720	113 -- 123	60 -- 70	-9.48758
0.0373636	0.94584	H375_680	94 -- 126	50 -- 85	-9.42535
0.0432702	0.94584	H375_8130	61 -- 89	70 -- 96	-9.17281
0.0450156	0.94584	H375_9060	126 -- 149	69 -- 97	-9.10438
0.0482611	0.94584	H375_8370	107 -- 149	66 -- 100	-8.9835

0.048511	0.94584	H375_2790	101 -- 131	68 -- 94	-8.97451
0.0486718	0.94584	H375_2850	103 -- 119	77 -- 93	-8.96875
0.0492938	0.94584	H375_7180	86 -- 98	63 -- 75	-8.94663
0.0504248	0.94584	H375_8280	51 -- 69	80 -- 95	-8.90707

*Rp*_sR67

TargetRNA2

Rank	Synonym	Energy	Pvalue	sRNA_start	sRNA_stop	mRNA_start	mRNA_stop
1	H375_2780	-11.58	0.01	77	85	-27	-19
2	H375_2910	-11.29	0.012	21	37	-6	10
3	H375_4170	-11.17	0.012	172	184	-64	-53
4	H375_980	-11.06	0.013	151	164	-71	-58
5	H375_5470	-10.94	0.014	96	113	-6	13
6	H375_8490	-10.64	0.017	153	169	6	20
7	H375_6610	-10.35	0.02	210	224	-7	7
8	H375_4810	-10.13	0.022	206	221	-71	-56
9	H375_7880	-10.12	0.022	1	16	-13	6
10	H375_8130	-10.11	0.022	6	23	2	19
11	H375_6770	-10.04	0.023	151	162	-31	-20
12	H375_7690	-9.98	0.024	154	165	-22	-11
13	H375_4660	-9.91	0.024	154	169	-44	-29
14	H375_9020	-9.6	0.028	136	148	-80	-68
15	H375_8420	-9.57	0.029	176	186	-76	-66
16	H375_8110	-9.55	0.029	176	190	-55	-41
17	H375_7440	-9.32	0.032	170	184	-80	-69
18	H375_1440	-9.22	0.034	216	230	-59	-45
19	H375_2540	-9.11	0.035	135	145	5	15
20	H375_8560	-9.03	0.037	204	215	-79	-68
21	H375_2550	-8.96	0.038	169	185	-2	15
22	H375_1110	-8.88	0.039	117	126	-77	-68
23	H375_5400	-8.75	0.041	58	72	-79	-65
24	H375_3710	-8.73	0.042	113	130	-77	-62
25	H375_5520	-8.69	0.042	177	187	-26	-16
26	H375_3730	-8.67	0.043	154	170	-61	-44
27	H375_5860	-8.62	0.044	128	141	-74	-61
28	H375_1350	-8.55	0.045	149	163	-38	-24
29	H375_8360	-8.53	0.045	192	204	-54	-42
30	H375_6090	-8.36	0.049	220	230	7	18

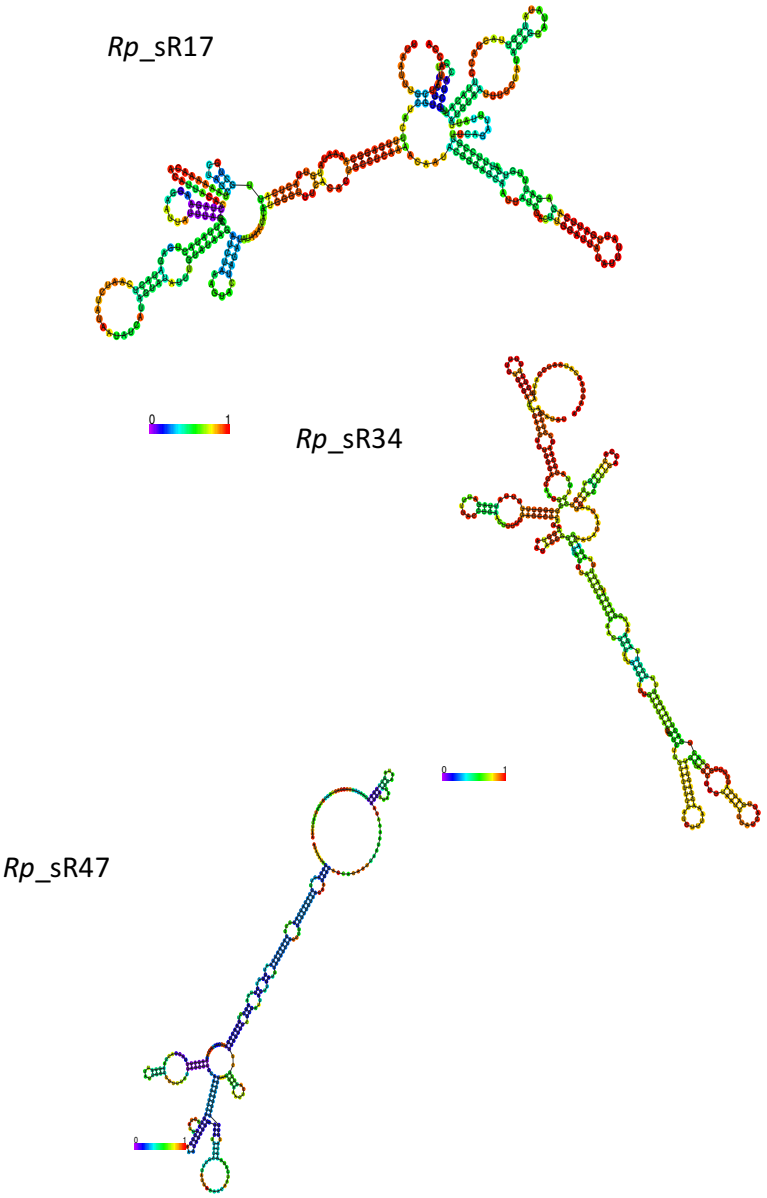
IntaRNA

p-value	fdr value	Target	Position	Position	Energy
0.000252	0.16965	H375_6400	111 -- 148	3 -- 44	-14.3456

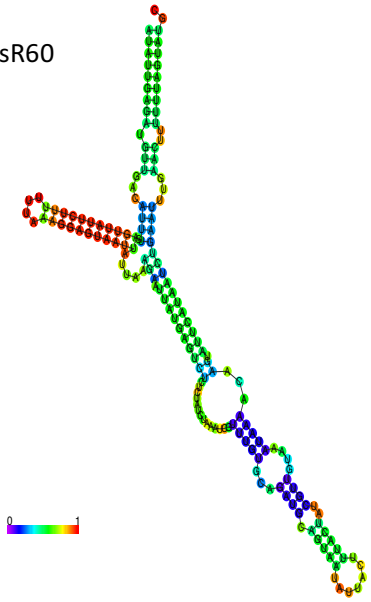
0.000369	0.16965	H375_6200	112 -- 122	119 -- 129	-13.9723
0.00196	0.32499	H375_8180	122 -- 136	18 -- 32	-12.238
0.001967	0.32499	H375_8870	5 -- 54	14 -- 66	-12.2343
0.002077	0.32499	H375_5660	3 -- 15	201 -- 213	-12.1752
0.002274	0.32499	H375_6530	41 -- 70	48 -- 75	-12.0757
0.002473	0.32499	H375_4890	111 -- 148	37 -- 72	-11.9831
0.003836	0.4411	H375_1080	134 -- 143	194 -- 203	-11.4913
0.005731	0.54303	H375_6100	88 -- 143	133 -- 174	-11.0299
0.005903	0.54303	H375_3950	42 -- 70	45 -- 74	-10.9955
0.008803	0.69301	H375_1120	2 -- 25	111 -- 130	-10.5238
0.009166	0.69301	H375_7880	64 -- 78	2 -- 15	-10.4754
0.010211	0.69301	H375_6690	6 -- 18	124 -- 136	-10.3456
0.010546	0.69301	H375_750	112 -- 149	42 -- 79	-10.3067
0.0135	0.77869	H375_5040	2 -- 28	194 -- 219	-10.0058
0.013543	0.77869	H375_2780	50 -- 56	78 -- 84	-10.0019
0.01612	0.82701	H375_4810	2 -- 25	119 -- 150	-9.78657
0.016181	0.82701	H375_1960	5 -- 45	138 -- 173	-9.78193
0.019375	0.86033	H375_3760	118 -- 134	1 -- 15	-9.55654
0.020038	0.86033	H375_3840	95 -- 134	28 -- 74	-9.51415
0.020128	0.86033	H375_7710	102 -- 122	113 -- 131	-9.50851
0.020573	0.86033	H375_5460	98 -- 111	204 -- 219	-9.48088
0.024148	0.92157	H375_680	23 -- 43	200 -- 217	-9.27723
0.024284	0.92157	H375_560	87 -- 137	3 -- 56	-9.27006
0.025862	0.92157	H375_1660	71 -- 91	127 -- 147	-9.18937
0.029352	0.92157	H375_5830	67 -- 87	44 -- 64	-9.02598
0.029354	0.92157	H375_1200	47 -- 74	149 -- 174	-9.0259
0.029978	0.92157	H375_540	111 -- 139	20 -- 49	-8.99858
0.030143	0.92157	H375_3710	2 -- 9	119 -- 126	-8.99148
0.032466	0.92157	H375_3900	31 -- 58	55 -- 82	-8.89472
0.033558	0.92157	H375_6090	139 -- 149	191 -- 201	-8.85145
0.034589	0.92157	H375_5110	60 -- 74	135 -- 148	-8.81176
0.038618	0.92157	H375_80	97 -- 108	119 -- 130	-8.66646
0.039296	0.92157	H375_1420	26 -- 50	146 -- 171	-8.64342
0.041053	0.92157	H375_940	1 -- 9	31 -- 39	-8.58528
0.041566	0.92157	H375_3910	108 -- 124	36 -- 53	-8.56876
0.042279	0.92157	H375_5580	3 -- 9	78 -- 84	-8.54607
0.042715	0.92157	H375_1700	64 -- 79	1 -- 15	-8.53238
0.042834	0.92157	H375_8760	120 -- 133	15 -- 29	-8.52867
0.044145	0.92157	H375_2890	98 -- 132	98 -- 135	-8.48834
0.044149	0.92157	H375_120	47 -- 67	108 -- 128	-8.48823
0.044663	0.92157	H375_1090	105 -- 116	102 -- 114	-8.47273

0.045763	0.92157	H375_5470	70 -- 87	97 -- 113	-8.44007
0.047108	0.92157	H375_600	6 -- 18	3 -- 15	-8.40112
0.048604	0.92157	H375_3590	22 -- 32	72 -- 82	-8.35898
0.049233	0.92157	H375_4330	64 -- 79	1 -- 16	-8.34161
0.049434	0.92157	H375_2550	80 -- 87	172 -- 179	-8.33612
0.049996	0.92157	H375_3670	134 -- 141	1 -- 8	-8.32083

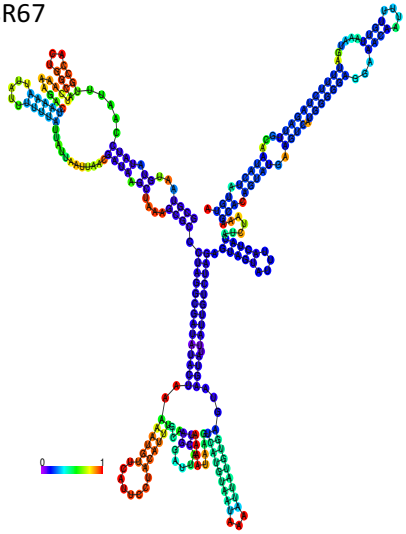
Additional File 5. Predicted secondary structure of novel sRNAs.
Predicted secondary structures for *Rp*_sR17, *Rp*_sR34, *Rp*_sR60, and *Rp*_sR67 as determined by RNA-fold. Color represents base-pairing probability from 0 to 1 (purple to red).



*Rp*_sR60



*Rp*_sR67



Additional File 6. *Rickettsia prowazekii* small RNAs (*Rp*_sR) identified in HMEC and AAE2.

RNA sequencing data from AAE2 and HMEC infected cell lines demonstrate that *R. prowazekii* strain Breinl encodes for at least 150 candidate sRNAs. Each of these candidates are listed here. The sRNA number, the approximate start location based on the start of the RNA sequencing reads, the approximate stop location based again on the RNA sequencing reads, the nucleotide size, the sRNA carrying strand, the type of sRNA, and its model host of infection are shown here. The column labeled as "Strand Orientation" refers to the orientation of the upstream gene, the trans-acting sRNA, and downstream gene. For cis-acting sRNAs, the corresponding ORF is listed in this column (F=forward; R=reverse; > = sense strand; < = anti-sense strand).

*Rp*_sR	Apporximate Start	Approximate Stop	Size (bp)	Strand	Strand Orientation	Type of sRNA	Model	Reference
1	10482	10278	204	R	</</<	Trans	HMEC	Schroeder, et al., 2015
2	10774	11003	230	F	</>/<	Trans	HMEC	Schroeder, et al., 2016
3	14215	13984	232	R	H375_160	Cis	HMEC	Schroeder, et al., 2016
4	15459	15203	257	R	H375_160	Cis	HMEC	Schroeder, et al., 2016
5	19850	19622	229	R	>/</>	Trans	HMEC	Schroeder, et al., 2016
6	22107	22298	192	F	</>/>	Trans	HMEC	Schroeder, et al., 2016
7	24935	25206	272	F	H375_230	Cis	HMEC	Schroeder, et al., 2016
8	47692	47542	150	F	>/>/>	Trans	HMEC	Schroeder, et al., 2015
9	48035	48200	166	F	>/>/>	Trans	HMEC	Schroeder, et al., 2016
10	48620	48834	215	F	>/>/>	Trans	HMEC	Schroeder, et al., 2016
11	71739	72043	305	F	H375_570	Cis	HMEC	Schroeder, et al., 2016
12	76189	76452	264	F	H375_570	Cis	HMEC	Schroeder, et al., 2016
13	77115	77335	221	F	</>/<	Trans	HMEC	Schroeder, et al., 2016
14	78259	78513	255	F	</>/>	Trans	HMEC	Schroeder, et al., 2016
15	88423	88616	194	F	>/>/<	Trans	HMEC	Schroeder, et al., 2016
16	100282	99698	585	R	H375_740	Cis	HMEC	Schroeder, et al., 2016
17	105748	105486	263	R	>/</>	Trans	HMEC	Schroeder, et al., 2016
18	116363	116200	164	R	375_870	Cis	HMEC	Schroeder, et al., 2016
19	132593	132781	189	F	</>/>	Trans	HMEC	Schroeder, et al., 2016
20	163641	163556	85	R	</</>	Trans	HMEC	Schroeder, et al., 2015
21	173793	173974	182	F	H375_1300	Cis	HMEC	Schroeder, et al., 2016

22	177789	178118	330	F	H375_1330	Cis	HMEC	Schroeder, et al., 2016
23	185267	185484	218	F	H375_1400	Cis	HMEC	Schroeder, et al., 2016
24	218432	218657	226	F	H375_1610	Cis	HMEC	Schroeder, et al., 2016
25	257587	257856	270	F	H375_1930	Cis	HMEC	Schroeder, et al., 2016
26	261134	261341	208	F	>/>/<	Trans	HMEC	Schroeder, et al., 2016
27	262322	262601	280	F	>/>/<	Trans	HMEC	Schroeder, et al., 2016
28	275336	275043	294	R	H375_2050	Cis	HMEC	Schroeder, et al., 2016
29	306070	305924	146	F	>/>/>	Trans	HMEC	Schroeder, et al., 2015
30	308324	308042	282	R	>/</<	Trans	HMEC	Schroeder, et al., 2015
31	308652	308848	197	F	>/>/<	Trans	HMEC	Schroeder, et al., 2016
32	309308	309564	257	F	H375_2350	Cis	HMEC	Schroeder, et al., 2016
33	315408	315772	365	F	375_2380	Cis	HMEC	Schroeder, et al., 2016
34	326339	326637	299	F	H375_2470	Cis	HMEC	Schroeder, et al., 2016
35	334531	334333	199	R	H375_2560	Cis	HMEC	Schroeder, et al., 2016
36	342505	342292	214	R	H375_2630	Cis	HMEC	Schroeder, et al., 2016
37	344054	343930	125	R	>/</<	Trans	HMEC	Schroeder, et al., 2016
38	366166	366242	77	F	>/>/>	Trans	HMEC	Schroeder, et al., 2016
39	371859	371506	353	R	>/</<	Trans	HMEC	Schroeder, et al., 2015
40	373780	373983	204	F	H375_2910	Cis	HMEC	Schroeder, et al., 2016
41	378222	378333	112	F	H375_2940	Cis	HMEC	Schroeder, et al., 2016
42	407851	407640	212	R	>/</<	Trans	HMEC	Schroeder, et al., 2016
43	427133	426907	227	R	>/</>	Trans	HMEC	Schroeder, et al., 2016
44	457001	456876	125	F	>/>/>	Trans	HMEC	Schroeder, et al., 2015
45	458370	458555	227	R	>/</>	Trans	HMEC	Schroeder, et al., 2016
46	462025	462293	268	F	H375_3670	Cis	HMEC	Schroeder, et al., 2016
47	481518	481833	316	F	H375_3890	Cis	HMEC	Schroeder, et al., 2016
48	494600	494387	214	R	</</<	Trans	HMEC	Schroeder, et al., 2016
49	504250	504506	256	F	>/>/>	Trans	HMEC	Schroeder, et al., 2016

50	514132	513893	240	R	</</<	Trans	HMEC	Schroeder, et al., 2016
51	531332	531187	146	R	>/</<	Trans	HMEC	Schroeder, et al., 2016
52	535456	535256	201	R	>/</<	Trans	HMEC	Schroeder, et al., 2016
53	558930	558719	212	R	>/</<	Trans	HMEC	Schroeder, et al., 2016
54	644329	644199	130	R	</</<	Trans	HMEC	Schroeder, et al., 2015
55	659164	659057	107	R	</</<	Trans	HMEC	Schroeder, et al., 2015
56	669521	669772	252	F	>/>/<	Trans	HMEC	Schroeder, et al., 2016
57	697106	696934	173	R	>/</<	Trans	HMEC	Schroeder, et al., 2016
58	812662	812412	251	R	</</<	Trans	HMEC	Schroeder, et al., 2016
59	830223	830416	194	F	>/>/<	Trans	HMEC	Schroeder, et al., 2016
60	844283	844606	324	F	>/>/>	Trans	HMEC	Schroeder, et al., 2016
61	859237	859436	200	F	>/>/>	Trans	HMEC	Schroeder, et al., 2016
62	909871	910281	411	F	H375_7470	Cis	HMEC	Schroeder, et al., 2016
63	925625	925929	305	F	</>/<	Trans	HMEC	Schroeder, et al., 2016
64	958095	957827	269	R	</</<	Trans	HMEC	Schroeder, et al., 2016
65	968393	968664	272	F	>./>/>	Trans	HMEC	Schroeder, et al., 2016
66	973818	973352	467	R	>/</>	Trans	HMEC	Schroeder, et al., 2016
67	998167	997927	240	R	>/</<	Trans	HMEC	Schroeder, et al., 2015
68	1039473	1039278	195	R	</</<	Trans	HMEC	Schroeder, et al., 2015
69	1046344	1046671	328	F	</>/<	Trans	HMEC	Schroeder, et al., 2016
70	1105018	1104959	59	F	>/>/>	Trans	HMEC	Schroeder, et al., 2015
71	34024	33655	369	R	</</<	Trans	AAE2	This Study
72	52729	53096	367	F	H375_0430	Cis	AAE2	This Study
73	118628	118261	367	R	H375_0880	Cis	AAE2	This Study
74	122578	122328	250	R	>/>/<	Trans	AAE2	This Study
75	127781	127541	240	R	>/</<	Trans	AAE2	This Study
76	189364	189678	314	F	</>/<	Trans	AAE2	This Study
77	222332	222551	219	F	H375_1620	Cis	AAE2	This Study

78	241272	241565	293	F	H375_1790	Cis	AAE2	This Study
79	267602	268133	531	F	H375_1990	Cis	AAE2	This Study
80	284574	285043	469	F	H375_2130	Cis	AAE2	This Study
81	286154	285747	407	R	H375_2140	Cis	AAE2	This Study
82	295757	296248	491	F	H375_2250	Cis	AAE2	This Study
83	361118	361612	494	F	</>/<	Trans	AAE2	This Study
84	391263	391442	179	F	H375_3070	Cis	AAE2	This Study
85	425033	424623	410	R	H375_3360	Cis	AAE2	This Study
86	427133	426907	227	R	>/</>	Trans	AAE2	This Study
87	477214	477426	212	F	H375_3850	Cis	AAE2	This Study
88	489258	488950	308	R	H375_3950	Cis	AAE2	This Study
89	500073	500490	417	F	H375_4040	Cis	AAE2	This Study
90	539542	539873	331	F	</>/<	Trans	AAE2	This Study
91	584699	584545	154	R	H375_4730	Cis	AAE2	This Study
92	595553	596109	556	F	H375_4840	Cis	AAE2	This Study
93	599387	599922	535	F	</>/<	Trans	AAE2	This Study
94	602516	603103	587	F	H375_4880	Cis	AAE2	This Study
95	622277	622493	216	F	H375_5060	Cis	AAE2	This Study
96	628340	628515	175	F	H375_5100	Cis	AAE2	This Study
97	630555	630851	296	F	H375_5130	Cis	AAE2	This Study
98	633733	634011	278	F	H375_5170	Cis	AAE2	This Study
99	634851	635100	249	F	</>/<	Trans	AAE2	This Study
100	636193	636573	380	F	H375_5200	Cis	AAE2	This Study
101	638310	638501	191	F	H375_5210	Cis	AAE2	This Study
102	655081	655278	197	F	H375_5270	Cis	AAE2	This Study
103	658089	658416	327	F	</>/<	Trans	AAE2	This Study
104	683020	682557	463	R	H375_5560	Cis	AAE2	This Study

105	697487	697795	308	F	H375_5680	Cis	AAE2	This Study
106	704302	704573	271	F	H375_5740	Cis	AAE2	This Study
107	705082	705369	287	F	H375_5760	Cis	AAE2	This Study
108	707777	708027	250	F	H375_5770	Cis	AAE2	This Study
109	730523	730271	252	R	H375_5960	Cis	AAE2	This Study
110	737433	737109	324	R	H375_6020	Cis	AAE2	This Study
111	738771	738453	318	R	>/</>	Trans	AAE2	This Study
112	739152	738831	321	R	>/</>	Trans	AAE2	This Study
113	765006	765327	321	F	H375_6260	Cis	AAE2	This Study
114	770496	769890	606	R	H375_6320	Cis	AAE2	This Study
115	778258	777992	266	R	H375_6400	Cis	AAE2	This Study
116	789170	789007	163	R	H375_6520	Cis	AAE2	This Study
117	803483	803249	234	R	H375_6630	Cis	AAE2	This Study
118	805321	804858	463	R	H375_6650	Cis	AAE2	This Study
119	807020	806844	176	R	H375_6670	Cis	AAE2	This Study
120	819723	820502	779	F	H375_6770	Cis	AAE2	This Study
121	830189	829961	228	R	H375_6840	Cis	AAE2	This Study
122	838741	839229	488	F	H375_6960	Cis	AAE2	This Study
123	843915	843653	262	R	>/</>	Trans	AAE2	This Study
124	849025	848559	466	R	H375_7030 H375_7040	Cis	AAE2	This Study
125	853045	852638	407	R	H375_7070	Cis	AAE2	This Study
126	855770	855961	191	F	</>/<	Trans	AAE2	This Study
127	860556	860192	364	R	H375_7160	Cis	AAE2	This Study
128	863840	862996	844	R	H375_7160	Cis	AAE2	This Study
129	864742	864313	429	R	>/>/>	Trans	AAE2	This Study
130	865797	866550	753	R	>/</>	Trans	AAE2	This Study

131	886013	886309	296	F	H375_7320	Cis	AAE2	This Study
132	889427	889686	259	F	H375_7360	Cis	AAE2	This Study
133	891055	891525	470	F	H375_7370	Cis	AAE2	This Study
134	904918	904686	232	R	H375_7450	Cis	AAE2	This Study
135	914970	914376	594	R	>/</>	Trans	AAE2	This Study
136	915777	915587	190	R	>/</>	Trans	AAE2	This Study
137	919393	919072	321	R	H375_7550	Cis	AAE2	This Study
138	945764	945294	470	R	H375_7800	Cis	AAE2	This Study
139	950181	950583	402	F	>/>/<	Trans	AAE2	This Study
140	954943	954625	318	R	H375_7880	Cis	AAE2	This Study
141	956931	957132	201	F	H375_7890	Cis	AAE2	This Study
142	972717	972412	305	R	>/</>	Trans	AAE2	This Study
143	976190	975674	516	R	H375_8030	Cis	AAE2	This Study
144	981480	981209	271	R	H375_8090	Cis	AAE2	This Study
145	982025	981793	232	R	H375_8100	Cis	AAE2	This Study
146	992543	992331	212	R	>/</>	Trans	AAE2	This Study
147	993107	992826	281	R	>/</>	Trans	AAE2	This Study
148	993723	993433	290	R	H375_8270	Cis	AAE2	This Study
149	995664	995482	182	R	H375_8270	Cis	AAE2	This Study
150	1001532	1001813	281	F	H375_8280	Cis	AAE2	This Study
151	1007564	1007770	206	F	H375_8330	Cis	AAE2	This Study
152	1010238	1010386	148	F	H375_8370	Cis	AAE2	This Study
153	1015702	1015344	358	R	>/</>	Trans	AAE2	This Study
154	1017292	1016581	711	R	H375_8410	Cis	AAE2	This Study
155	1018205	1017776	429	R	>/</<	Trans	AAE2	This Study
156	1022112	1022334	222	F	H375_8450	Cis	AAE2	This Study
157	1038622	1038869	247	F	</>/<	Trans	AAE2	This Study

158	1039987	1040438	451	F	H375_8570	Cis	AAE2	This Study
159	1046344	1046671	328	F	</>/<	Trans	AAE2	This Study
160	1059802	1059493	309	R	H375_8730	Cis	AAE2	This Study
161	1062616	1062298	318	R	H375_8790	Cis	AAE2	This Study
162	1069032	1068730	302	R	H375_8930	Cis	AAE2	This Study
163	1070608	1070315	293	R	H375_8950	Cis	AAE2	This Study

Additional File 7. CopraRNA predicted targets for *Rp*_sR60.

The complete list of predicted targets by CopraRNA for *Rp*_sR60. Only those targets with a P≤0.05 were included for analysis.

Rank	CopraRNA p-value	Locus Tag	Annotation
1*	0.001253	h375_8740	50S ribosomal protein L3
2*	0.002567	h375_7670	uncharacterized protein
3*	0.00319	h375_1650	2-hydroxy-6-oxo-6-phenylhexa-2 4-dienoate hydrolase
4*	0.003375	h375_2860	DNA topoisomerase 1
5*	0.003515	h375_7680	uncharacterized protein
6*	0.004535	h375_1040	tRNA dimethylallyltransferase
7*	0.005017	h375_6080	Uncharacterized protein
8*	0.006133	h375_8760	50S ribosomal protein L23
9*	0.006635	h375_420	Chaperone protein HscA
10*	0.006775	h375_2090	Soluble lytic murein transglycosylase
11	0.007708	h375_9210	sensor histidine kinase NtrY-like protein
12	0.01089	h375_8890	50S ribosomal protein L18
13	0.01155	h375_370	Holo-[acyl-carrier-protein] synthase
14	0.01169	h375_8110	glutamine amidotransferase-like protein
15	0.01178	h375_7090	Maf-like protein
16	0.01262	h375_1430	Porphobilinogen deaminase
17	0.01521	h375_4920	50S ribosomal protein L7/L12
18	0.01612	h375_2890	uncharacterized protein
19	0.0177	h375_3140	Thymidylate synthase ThyX
20	0.01772	h375_8880	30S ribosomal protein S8
21	0.01777	h375_250	Protein MurJ
22	0.01783	h375_770	reductase
23	0.01794	h375_2780	UDP-glucose 4-epimerase
24	0.01867	h375_800	Uncharacterized protein
25	0.01903	h375_5790	Uncharacterized protein
26	0.01917	h375_4980	Elongation factor G
27	0.01967	h375_120	transporter
28	0.02008	h375_2730	Glycosyltransferase
29	0.02223	h375_2710	glycosyltransferase
30	0.02275	h375_8020	DNA ligase

31	0.02611	h375_3690	UDP-N-acetylenolpyruvoylglucosamine reductase
32	0.02672	h375_5610	Deoxycytidine triphosphate deaminase
33	0.02725	h375_2510	Uncharacterized protein
34	0.02731	h375_3820	Regulatory component of sensory transduction system
35	0.02824	h375_7620	reductase
36	0.03014	h375_830	Predicted membrane protein COG5346
37	0.03149	h375_4430	2-oxoglutarate dehydrogenase E1 component
38	0.03152	h375_4620	Ribosome association toxin RatA
39	0.03153	h375_3450	Cytochrome b
40	0.03185	h375_7880	Poly-beta-hydroxybutyrate polymerase
41	0.03302	h375_6920	Opacity protein and surface antigens COG3637
42	0.03347	h375_480	Cell division protein FtsL
43	0.03348	h375_1850	Phosphatidate cytidylyltransferase
44	0.04017	h375_2430	ABC-type uncharacterized transport system permease component
45	0.04196	h375_3420	Translation factor GUF1
46	0.04197	h375_2330	Malate dehydrogenase
47	0.04339	h375_4840	uncharacterized protein
48	0.04418	h375_6370	50S ribosomal protein L33
49	0.04494	h375_1570	Isopentenyl-diphosphate delta-isomerase
50	0.04751	h375_2640	Outer membrane assembly protein
51	0.04761	h375_3760	CDP-diacylglycerol--serine O-phosphatidyltransferase
52	0.04982	h375_3000	HlyD family secretion protein
53	0.04983	h375_3230	Type IV secretion system protein VirD4

Bibliography

Albrecht, M., Sharma, C.M., Reinhardt, R., Vogel, J. and Rudel, T. (2010) Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Res.* **38**: 868-877.

Alexeyev, M.F. and Winkler, H.H. (1999) Membrane topology of the *Rickettsia prowazekii* ATP/ADP translocase revealed by novel dual pho-lac reporters. *J. Mol. Biol.* **285**: 1503-1513.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.

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

Ammerman, N.C., Gillespie, J.J., Neuwald, A.F., Sobral, B.W. and Azad, A.F. (2009) A typhus group-specific protease defies reductive evolution in rickettsiae. *J. Bacteriol.* **191**: 7609-7613.

Andersson, S.G.E., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C.M. and Podowski, R.M. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**: 133-143.

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H.*, et al.* (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* **11**: 941-950.

Audia, J.P., Patton, M.C. and Winkler, H.H. (2008) DNA microarray analysis of the heat shock transcriptome of the obligate intracytoplasmic pathogen *Rickettsia prowazekii*. *Appl. Environ. Microbiol.* **74**: 7809-7812.

Azad, A.F. and Beard, C.B. (1998) Rickettsial pathogens and their arthropod vectors. *Emerging Infect. Dis.* **4**: 179-186.

Azad, A.F., Sacci, J.B., Nelson, W.M., Dasch, G.A., Schmidtmann, E.T. and Carl, M. (1992) Genetic characterization and transovarial transmission of a typhus-like *Rickettsia* found in cat fleas. *Proc. Natl. Acad. Sci. U. S. A.* **89**: 43-46.

Badiaga, S. and Brouqui, P. (2012) Human louse-transmitted infectious diseases. *Clin. Microbiol. Infect.* **18**: 332-337.

Bandara, A.B., Sriranganathan, N., Schurig, G.G. and Boyle, S.M. (2005) Carboxyl-terminal protease regulates *Brucella suis* morphology in culture and persistence in macrophages and mice. *J. Bacteriol.* **187**: 5767-5775.

Barrick, J.E., Sudarsan, N., Weinberg, Z., Ruzzo, W.L. and Breaker, R.R. (2005) 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *RNA* **11**: 774-784.

Basu, D., Khare, G., Singh, S., Tyagi, A., Khosla, S. and Mande, S.C. (2009) A novel nucleoid-associated protein of *Mycobacterium tuberculosis* is a sequence homolog of GroEL. *Nucleic Acids Res.* **37**: 4944-4954.

Bechah, Y., Capo, C., Mege, J.L. and Raoult, D. (2008a) Epidemic typhus. *Lancet Infect. Dis.* **8**: 417-426.

Bechah, Y., Capo, C., Mege, J.L. and Raoult, D. (2008b) Rickettsial diseases: from *Rickettsia*-arthropod relationships to pathophysiology and animal models. *Future Microbiol.* **3**: 223-236.

Blanc, G., Ogata, H., Robert, C., Audic, S., Suhre, K., Vestris, G.*, et al.* (2007) Reductive genome evolution from the mother of *Rickettsia*. *PLoS Genet.* **3**: e14.

Bobrovskyy, M., Vanderpool, C.K. and Richards, G.R. (2015) Small RNAs regulate primary and secondary metabolism in Gram-negative bacteria. *Microbiology Spectrum* **3**.

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M.*, et al.* (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* **21**: 1353-1366.

Bozeman, F.M., Masiello, S.A., Williams, M.S. and Elisberg, B.L. (1975) Epidemic typhus rickettsiae isolated from flying squirrels. *Nature* **255**: 545-547.

Busch, A., Richter, A.S. and Backofen, R. (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**: 2849-2856.

Caldelari, I., Chao, Y., Romby, P. and Vogel, J. (2013) RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb. Perspect. Med.* **3**: a010298.

Chabelskaya, S., Gaillot, O. and Felden, B. (2010) A *Staphylococcus aureus* small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS Path.* **6**: e1000927.

Chan, Y.G., Cardwell, M.M., Hermanas, T.M., Uchiyama, T. and Martinez, J.J. (2009) Rickettsial outer-membrane protein B (rOmpB) mediates bacterial invasion through Ku70 in an actin, c-Cbl, clathrin and caveolin 2-dependent manner. *Cell. Microbiol.* **11**: 629-644.

Chao, Y. and Vogel, J. (2010) The role of Hfq in bacterial pathogens. *Curr. Opin. Microbiol.* **13**: 24-33.

Childs, J.E. and Paddock, C.D. (2002) Passive surveillance as an instrument to identify risk factors for fatal Rocky Mountain spotted fever: is there more to learn? *Am. J. Trop. Med. Hyg.* **66**: 450-457.

Clark, T.R., Noriea, N.F., Bublitz, D.C., Ellison, D.W., Martens, C., Lutter, E.I.*, et al.* (2015) Comparative genome sequencing of *Rickettsia rickettsii* strains that differ in virulence. *Infect. Immun.* **83**: 1568-1576.

Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**: 1188-1190.

Darby, A.C., Cho, N.H., Fuxelius, H.H., Westberg, J. and Andersson, S.G. (2007) Intracellular pathogens go extreme: genome evolution in the Rickettsiales. *Trends Genet.* **23**: 511-520.

DebRoy, S., Gebbie, M., Ramesh, A., Goodson, J.R., Cruz, M.R., van Hoof, A.*, et al.* (2014) A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. *Science* **345**: 937-940.

DiChiara, J.M., Contreras-Martinez, L.M., Livny, J., Smith, D., McDonough, K.A. and Belfort, M. (2010) Multiple small RNAs identified in *Mycobacterium bovis* BCG are also expressed in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. *Nucleic Acids Res.* **38**: 4067-4078.

Drexler, N.A., Dahlgren, F.S., Heitman, K.N., Massung, R.F., Paddock, C.D. and Behravesh, C.B. (2016) National surveillance of spotted fever group Rickettsioses in the United States, 2008-2012. *Am. J. Trop. Med. Hyg.* **94**: 26-34.

Driskell, L.O., Yu, X.J., Zhang, L., Liu, Y., Popov, V.L., Walker, D.H.*, et al.* (2009) Directed mutagenesis of the *Rickettsia prowazekii* pld gene encoding phospholipase D. *Infect. Immun.* **77**: 3244-3248.

Dugar, G., Herbig, A., Förstner, K.U., Heidrich, N., Reinhardt, R., Nieselt, K.*, et al.* (2013) High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* **9**: e1003495.

El Karkouri, K., Pontarotti, P., Raoult, D. and Fournier, P.-E. (2016) Origin and evolution of Rickettsial plasmids. *PLoS One* **11**: e0147492.

Ellison, D.W., Clark, T.R., Sturdevant, D.E., Virtaneva, K. and Hackstadt, T. (2009) Limited transcriptional responses of *Rickettsia rickettsii* exposed to environmental stimuli. *PLoS One* **4**: e5612.

Ellison, D.W., Clark, T.R., Sturdevant, D.E., Virtaneva, K., Porcella, S.F. and Hackstadt, T. (2008) Genomic comparison of virulent *Rickettsia rickettsii* Sheila Smith and avirulent *Rickettsia rickettsii* Iowa. *Infect. Immun.* **76**: 542-550.

Eremeeva, M.E., Dasch, G.A. and Silverman, D.J. (2003) Evaluation of a PCR assay for quantitation of *Rickettsia rickettsii* and closely related spotted fever group Rickettsiae. *J. Clin. Microbiol.* **41**: 5466-5472.

Farhang-Azad, A., Traub, R. and Baqar, S. (1985) Transovarial transmission of murine typhus rickettsiae in *Xenopsylla cheopis* fleas. *Science* **227**: 543-545.

Faucher, S.P., Friedlander, G., Livny, J., Margalit, H. and Shuman, H.A. (2010) *Legionella pneumophila* 6S RNA optimizes intracellular multiplication. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 7533-7538.

Feng, L., Rutherford, Steven T., Papenfort, K., Bagert, John D., van Kessel, Julia C., Tirrell, David A.*, et al.* (2015) A Qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* **160**: 228-240.

Fournier, P.E., El Karkouri, K., Leroy, Q., Robert, C., Giumelli, B., Renesto, P.*, et al.* (2009) Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction. *BMC Genomics* **10**: 166.

Fujita, H., Fournier, P.E., Takada, N., Saito, T. and Raoult, D. (2006) *Rickettsia asiatica* sp. nov., isolated in Japan. *Int. J. Syst. Evol. Microbiol.* **56**: 2365-2368.

Fuxelius, H.-H., Darby, A., Min, C.-K., Cho, N.-H. and Andersson, S.G.E. (2007) The genomic and metabolic diversity of *Rickettsia*. *Res. Microbiol.* **158**: 745-753.

Gautheret, D. and Lambert, A. (2001) Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J. Mol. Biol.* **313**: 1003-1011.

Georgiades, K., Merhej, V., El Karkouri, K., Raoult, D. and Pontarotti, P. (2011) Gene gain and loss events in *Rickettsia* and *Orientia* species. *Biol. Direct* **6**: 6.

Georgiades, K. and Raoult, D. (2011) Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. *PLoS One* **6**: e17962.

Gillespie, J.J., Ammerman, N.C., Dreher-Lesnick, S.M., Rahman, M.S., Worley, M.J., Setubal, J.C., *et al.* (2009) An anomalous type IV secretion system in *Rickettsia* is evolutionarily conserved. *PLoS One* **4**: e4833.

Gillespie, J.J., Beier, M.S., Rahman, M.S., Ammerman, N.C., Shallom, J.M., Purkayastha, A., *et al.* (2007) Plasmids and rickettsial evolution: insight from *Rickettsia felis*. *PLoS One* **2**: e266.

Gillespie, J.J., Joardar, V., Williams, K.P., Driscoll, T., Hostetler, J.B., Nordberg, E., *et al.* (2012) A *Rickettsia* genome overrun by mobile genetic elements provides insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. *J. Bacteriol.* **194**: 376-394.

Gillespie, J.J., Kaur, S.J., Rahman, M.S., Rennoll-Bankert, K., Sears, K.T., Beier-Sexton, M., *et al.* (2015) Secretome of obligate intracellular *Rickettsia*. *FEMS Microbiol. Rev.* **39**: 47-80.

Gillespie, J.J., Williams, K., Shukla, M., Snyder, E.E., Nordberg, E.K., Ceraul, S.M., *et al.* (2008) *Rickettsia* phylogenomics: unwinding the intricacies of obligate intracellular life. *PLoS One* **3**: e2018.

Gómez-Lozano, M., Marvig, R.L., Tulstrup, M.V. and Molin, S. (2014) Expression of antisense small RNAs in response to stress in *Pseudomonas aeruginosa*. *BMC Genomics* **15**: 1-11.

Gong, H., Vu, G.P., Bai, Y., Chan, E., Wu, R., Yang, E., *et al.* (2011) A *Salmonella* small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. *PLoS Path.* **7**: e1002120.

Gottesman, S. and Storz, G. (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* **3**: a003798-a003798.

Grieshaber, N.A., Grieshaber, S.S., Fischer, E.R. and Hackstadt, T. (2006) A small RNA inhibits translation of the histone-like protein Hc1 in *Chlamydia trachomatis*. *Mol. Microbiol.* **59**: 541-550.

Gripenland, J., Netterling, S., Loh, E., Tiensuu, T., Toledo-Arana, A. and Johansson, J. (2010) RNAs: regulators of bacterial virulence. *Nat. Rev. Microbiol.* **8**: 857-866.

Guillier, M. and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.* **59**: 231-247.

Hansen, A.K. and Degnan, P.H. (2014) Widespread expression of conserved small RNAs in small symbiont genomes. *ISME Journal* **8**: 2490-2502.

Harley, C.B. and Reynolds, R.P. (1987) Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**: 2343-2361.

Hindley, J. (1967) Fractionation of 32p-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. *J. Mol. Biol.* **30**: 125-&.

Hofacker, I.L., (2002) RNA secondary structure analysis using the Vienna RNA package. In: Current Protocols in Bioinformatics. John Wiley & Sons, Inc., pp.

Hoge, R., Laschinski, M., Jaeger, K.E., Wilhelm, S. and Rosenau, F. (2011) The subcellular localization of a C-terminal processing protease in *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* **316**: 23-30.

Holste, D., Weiss, O., Grosse, I. and Herzel, H. (2000) Are noncoding sequences of *Rickettsia prowazekii* remnants of "neutralized" genes? *J. Mol. Biol.* **51**: 353-362.

Huerta, A.M. and Collado-Vides, J. (2003) Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. *J. Mol. Biol.* **333**: 261-278.

Janion, C. (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int. J. Biol. Sci.* **4**: 338-344.

Khandige, S., Kronborg, T., Uhlin, B.E. and Møller-Jensen, J. (2015) sRNA-mediated regulation of P-fimbriae phase variation in uropathogenic *Escherichia coli*. *PLoS Path.* **11**: e1005109.

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

Kroger, C., Dillon, S.C., Cameron, A.D., Papenfort, K., Sivasankaran, S.K., Hokamp, K.*, et al.* (2012) The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc. Natl. Acad. Sci. U. S. A.* **109**: E1277-1286.

Lad, S.P., Yang, G., Scott, D.A., Wang, G., Nair, P., Mathison, J.*, et al.* (2007) Chlamydial CT441 is a PDZ domain-containing tail-specific protease that interferes with the NF-κB pathway of immune response. *J. Bacteriol.* **189**: 6619-6625.

Lafri, I., Leulmi, H., Baziz-Neffah, F., Lalout, R., Mohamed, C., Mohamed, K.*, et al.* (2015) Detection of a novel *Rickettsia* sp. in soft ticks (Acari: Argasidae) in Algeria. *Microb. Infect.* **17**: 859-861.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357-359.

Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., de los Mozos, I.R., Vergara-Irigaray, M.*, et al.* (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 20172-20177.

Lasa, I., Toledo-Arana, A. and Gingeras, T.R. (2012) An effort to make sense of antisense transcription in bacteria. *RNA Biol.* **9**: 1039-1044.

Lawley, T.D., Klimke, W.A., Gubbins, M.J. and Frost, L.S. (2003) F factor conjugation is a true type IV secretion system. *FEMS Microbiol. Lett.* **224**: 1-15.

Lee, E.J. and Groisman, E.A. (2010) An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol. Microbiol.* **76**: 1020-1033.

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

Leroy, Q., Lebrigand, K., Armougom, F., Barbry, P., Thiery, R. and Raoult, D. (2010) *Coxiella burnetii* transcriptional analysis reveals serendipity clusters of regulation in intracellular bacteria. *PLoS One* **5**: e15321.

Li, L., Huang, D., Cheung, M.K., Nong, W., Huang, Q. and Kwan, H.S. (2013) BSRD: a repository for bacterial small regulatory RNA. *Nucleic Acids Res.* **41**: D233-D238.

Liu, J.M. and Camilli, A. (2010) A broadening world of bacterial small RNAs. *Curr. Opin. Microbiol.* **13**: 18-23.

Liu, X., Brutlag, D.L. and Liu, J.S. (2001) BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac. Symp. Biocomput.*: 127-138.

Livny, J., Brencic, A., Lory, S. and Waldor, M.K. (2006) Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. *Nucleic Acids Res.* **34**: 3484-3493.

Livny, J., Teonadi, H., Livny, M. and Waldor, M.K. (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS One* **3**: e3197.

Lu, X., Goodrich-Blair, H. and Tjaden, B. (2011) Assessing computational tools for the discovery of small RNA genes in bacteria. *RNA* **17**: 1635-1647.

Macke, T.J., Ecker, D.J., Gutell, R.R., Gautheret, D., Case, D.A. and Sampath, R. (2001) RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Res.* **29**: 4724-4735.

Masse, E. and Gottesman, S. (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 4620-4625.

Mathews, D.H., Moss, W.N. and Turner, D.H. (2010) Folding and finding RNA secondary structure. *Cold Spring Harb. Perspect. Biol.* **2**: a003665.

McLeod, M.P., Qin, X., Karpathy, S.E., Gioia, J., Highlander, S.K., Fox, G.E.*, et al.* (2004) Complete genome sequence of *Rickettsia typhi* and comparison with sequences of other rickettsiae. *J. Bacteriol.* **186**: 5842-5855.

Medina-Sanchez, A., Bouyer, D.H., Alcantara-Rodriguez, V., Mafra, C., Zavala-Castro, J., Whitworth, T.*, et al.* (2005) Detection of a typhus group *Rickettsia* in *Amblyomma* ticks in the state of Nuevo Leon, Mexico. *Ann. N. Y. Acad. Sci.* **1063**: 327-332.

Mellin, J.R., Koutero, M., Dar, D., Nahori, M.-A., Sorek, R. and Cossart, P. (2014) Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. *Science* **345**: 940-943.

Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B.*, et al.* (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One* **4**: e7526.

Merhej, V., Angelakis, E., Socolovschi, C. and Raoult, D. (2014) Genotyping, evolution and epidemiological findings of *Rickettsia* species. *Infect., Genet. Evol.* **25**: 122-137.

Merhej, V., Notredame, C., Royer-Carenzi, M., Pontarotti, P. and Raoult, D. (2011) The rhizome of life: the sympatric *Rickettsia felis* paradigm demonstrates the random transfer of DNA sequences. *Mol. Biol. Evol.* **28**: 3213-3223.

Merhej, V. and Raoult, D. (2011) Rickettsial evolution in the light of comparative genomics. *Biological Reviews* **86**: 379-405.

Mitchell, J.E., Zheng, D.L., Busby, S.J.W. and Minchin, S.D. (2003) Identification and analysis of 'extended-10' promoters in *Escherichia coli*. *Nucleic Acids Res.* **31**: 4689-4695.

Moody, M.J., Young, R.A., Jones, S.E. and Elliot, M.A. (2013) Comparative analysis of non-coding RNAs in the antibiotic-producing *Streptomyces* bacteria. *BMC Genomics* **14**: 558.

Morita, T., Maki, K. and Aiba, H. (2012) Detection of sRNA-mRNA interactions by electrophoretic mobility shift assay. *Methods Mol. Biol.* **905**: 235-244.

Mraheil, M.A., Billion, A., Mohamed, W., Mukherjee, K., Kuenne, C., Pischimarov, J.*, et al.* (2011) The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. *Nucleic Acids Res.* **39**: 4235-4248.

Munderloh, U.G. and Kurtti, T.J. (1989) Formulation of medium for tick cell culture. *Exp. Appl. Acarol.* **7**: 219-229.

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H.E., Moran, N.A.*, et al.* (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314**: 267-267.

Nielsen, J.S., Larsen, M.H., Lillebaek, E.M., Bergholz, T.M., Christiansen, M.H., Boor, K.J.*, et al.* (2011) A small RNA controls expression of the chitinase ChiA in *Listeria monocytogenes*. *PLoS One* **6**: e19019.

Ogata, H., Audic, S., Renesto-Audiffren, P., Fournier, P.E., Barbe, V., Samson, D.*, et al.* (2001) Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293**: 2093-2098.

Ogata, H., Renesto, P., Audic, S., Robert, C., Blanc, G., Fournier, P.E.*, et al.* (2005) The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. *PLoS Biol.* **3**: e248.

Ogata, H., Scola, B., Audic, S., Renesto, P., Blanc, G. and Robert, C. (2006) Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet.* **2**.

Oliva, G., Sahr, T. and Buchrieser, C. (2015) Small RNAs, 5' UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence. *FEMS Microbiol. Rev.* **39**: 331-349.

Ooi, W.F., Ong, C., Nandi, T., Kreisberg, J.F., Chua, H.H., Sun, G.*, et al.* (2013) The condition-dependent transcriptional landscape of *Burkholderia pseudomallei*. *PLoS Genet.* **9**: e1003795.

Openshaw, J.J., Swerdlow, D.L., Krebs, J.W., Holman, R.C., Mandel, E., Harvey, A.*, et al.* (2010) Rocky Mountain spotted fever in the United States, 2000–2007: Interpreting contemporary increases in incidence. *Am. J. Trop. Med. Hyg.* **83**: 174-182.

Östberg, Y., Bunikis, I., Bergström, S. and Johansson, J. (2004) The etiological agent of Lyme Disease, *Borrelia burgdorferi*, appears to contain only a few small RNA molecules. *J. Bacteriol.* **186**: 8472-8477.

Otaka, H., Ishikawa, H., Morita, T. and Aiba, H. (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 13059-13064.

Ozsolak, F. and Milos, P.M. (2011) RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* **12**: 87-98.

Padalon-Brauch, G., Hershberg, R., Elgrably-Weiss, M., Baruch, K., Rosenshine, I., Margalit, H.*, et al.* (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res.* **36**: 1913-1927.

Pain, A., Ott, A., Amine, H., Rochat, T., Bouloc, P. and Gautheret, D. (2015) An assessment of bacterial small RNA target prediction programs. *RNA Biol.* **12**: 509-513.

Papenfort, K. and Vanderpool, C.K. (2015) Target activation by regulatory RNAs in bacteria. *FEMS Microbiol. Rev.* **39**: 362-378.
Parola, P., Paddock, C.D. and Raoult, D. (2005) Tick-borne rickettsioses around the world: emerging diseases challenging old concepts. *Clin. Microbiol. Rev.* **18**.
Parulekar, R.S., Barage, S.H., Jalkute, C.B., Dhanavade, M.J., Fandilolu, P.M. and Sonawane, K.D. (2013) Homology modeling, molecular docking and DNA binding studies of nucleotide excision repair UvrC protein from *M. tuberculosis*. *Protein J.* **32**: 467-476.
Philip, C.B. (1959) Some epidemiological considerations in Rocky Mountain spotted fever. *Public Health Rep.* **74**: 595-600.
Philip, C.B., Hoogstraal, H., Reiss-Gutfreund, R. and Clifford, C.M. (1966) Evidence of rickettsial disease agents in ticks from Ethiopian cattle. *Bull. W.H.O.* **35**: 127-131.
Pichon, C. and Felden, B. (2008) Small RNA gene identification and mRNA target predictions in bacteria. *Bioinformatics* **24**: 2807-2813.
Raghavan, R., Groisman, E.A. and Ochman, H. (2011) Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.* **21**: 1487-1497.
Raghavan, R., Kacharia, F.R., Millar, J.A., Sislak, C.D. and Ochman, H. (2015) Genome rearrangements can make and break small RNA genes. *Genome Biol. Evol.* **7**: 557-566.
Raghavan, R., Sloan, D.B. and Ochman, H. (2012) Antisense transcription is pervasive but rarely conserved in enteric bacteria. *mBio* **3**.
Raoult, D., Dutour, O., Houhamdi, L., Jankauskas, R., Fournier, P.-E., Ardagna, Y.*, et al.* (2006) Evidence for louse-transmitted diseases in soldiers of Napoleon's Grand Army in Vilnius. *J. Infect. Dis.* **193**: 112-120.
Raoult, D., Woodward, T. and Dumler, J.S. (2004) The history of epidemic typhus. *Infect. Dis. Clin. North Am.* **18**: 127-140.
Reeves, W.K., Murray, K.O., Meyer, T.E., Bull, L.M., Pascua, R.F., Holmes, K.C.*, et al.* (2008) Serological evidence of typhus group *Rickettsia* in a homeless population in Houston, Texas. *J. Vector Ecol.* **33**: 205-207.
Reiss-Gutfreund, R.J. (1966) The isolation of *Rickettsia prowazeki* and *mooseri* from unusual sources. *Am. J. Trop. Med. Hyg.* **15**: 943-949.
Renesto, P., Rovery, C., Schrenzel, J., Leroy, Q., Huyghe, A., Li, W.*, et al.* (2008) *Rickettsia conorii* transcriptional response within inoculation eschar. *PLoS One* **3**: e3681.
Rieder, R., Reinhardt, R., Sharma, C. and Vogel, J. (2012) Experimental tools to identify RNA-protein interactions in *Helicobacter pylori*. *RNA Biol.* **9**: 520-531.
Rivas, E. and Eddy, S.R. (2001) Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics* **2**: 8.
Rogozin, I.B., Makarova, K.S., Natale, D.A., Spiridonov, A.N., Tatusov, R.L., Wolf, Y.I.*, et al.* (2002) Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. *Nucleic Acids Res.* **30**: 4264-4271.
Roux, V., Rydkina, E., Eremeeva, M. and Raoult, D. (1997) Citrate synthase gene comparison, a new tool for phylogenetic analysis, and its application for the rickettsiae. *Int. J. Syst. Bacteriol.* **47**: 252-261.

Rydkina, E., Sahni, A., Silverman, D.J. and Sahni, S.K. (2007) Comparative analysis of host-cell signalling mechanisms activated in response to infection with *Rickettsia conorii* and *Rickettsia typhi*. *J. Med. Microbiol.* **56**: 896-906.

Rydkina, E., Silverman, D.J. and Sahni, S.K. (2005) Activation of p38 stress-activated protein kinase during *Rickettsia rickettsii* infection of human endothelial cells: role in the induction of chemokine response. *Cell. Microbiol.* **7**: 1519-1530.

Rydkina, E., Turpin, L.C. and Sahni, S.K. (2010) *Rickettsia rickettsii* infection of human macrovascular and microvascular endothelial cells reveals activation of both common and cell type-specific host response mechanisms. *Infect. Immun.* **78**: 2599-2606.

Sahni, S.K., Narra, H.P., Sahni, A. and Walker, D.H. (2013) Recent molecular insights into rickettsial pathogenesis and immunity. *Future Microbiol.* **8**: 1265-1288.

Sahni, S.K., Rydkina, E., Joshi, S.G., Sporn, L.A. and Silverman, D.J. (2003) Interactions of *Rickettsia rickettsii* with endothelial Nuclear Factor-κB in a "cell-free" system. *Ann. N. Y. Acad. Sci.* **990**: 635-641.

Sahni, S.K., Van Antwerp, D.J., Eremeeva, M.E., Silverman, D.J., Marder, V.J. and Sporn, L.A. (1998) Proteasome-independent activation of nuclear factor kappaB in cytoplasmic extracts from human endothelial cells by *Rickettsia rickettsii*. *Infect. Immun.* **66**: 1827-1833.

Schroeder, C.L., Narra, H.P., Rojas, M., Sahni, A., Patel, J., Khanipov, K.*, et al.* (2015) Bacterial small RNAs in the Genus *Rickettsia*. *BMC Genomics* **16**: 1075.

Schroeder, C.L.C., Narra, H.P., Sahni, A., Rojas, M., Khanipov, K., Patel, J.*, et al.* (2016) Identification and characterization of novel small RNAs in *Rickettsia prowazekii*. *Front. Microbiol.* **7**.

Seo, J. and Darwin, A.J. (2013) The *Pseudomonas aeruginosa* periplasmic protease CtpA can affect systems that impact its ability to mount both acute and chronic infections. *Infect. Immun.* **81**: 4561-4570.

Sharma, C. and Heidrich, N. (2012) Small RNAs and virulence in bacterial pathogens. *RNA Biol.* **9**: 361-363.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A.*, et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**: 250-255.

Sharma, C.M. and Vogel, J. (2014) Differential RNA-seq: the approach behind and the biological insight gained. *Curr. Opin. Microbiol.* **19**: 97-105.

Sharma, R., Arya, S., Patil, S.D., Sharma, A., Jain, P.K., Navani, N.K.*, et al.* (2014) Identification of novel regulatory small RNAs in *Acinetobacter baumannii*. *PLoS One* **9**: e93833.

Shinhara, A., Matsui, M., Hiraoka, K., Nomura, W., Hirano, R., Nakahigashi, K.*, et al.* (2011) Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC Genomics* **12**: 428.

Smadel, J.E. (1959) Status of the Rickettsioses in the United States. *Ann. Intern. Med.* **51**: 421-435.

Socolovschi, C., Mediannikov, O., Raoult, D. and Parola, P. (2009) The relationship between spotted fever group *Rickettsiae* and Ixodid ticks. *Vet. Res.* **40**: 34.

Solovyey, V. and Salamov, A., (2010) Automatic annotation of microbial genomes and metagenomic sequences. In: Metagenomics and its Applications in Agriculture,

Biomedicine and Environmental Studies. R.W. Li (ed). Nova Science Publishers, pp. 61-78.
Soucy, S.M., Huang, J. and Gogarten, J.P. (2015) Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* **16**: 472-482.
Sridhar, J., Sambaturu, N., Sabarinathan, R., Ou, H.Y., Deng, Z., Sekar, K.*, et al.* (2010) sRNAscanner: a computational tool for intergenic small RNA detection in bacterial genomes. *PLoS One* **5**: e11970.
Stazic, D. and Voss, B. (2016) The complexity of bacterial transcriptomes. *J. Biotechnol.* **232**: 69-78.
Stubben, C.J., Micheva-Viteva, S.N., Shou, Y., Buddenborg, S.K., Dunbar, J.M. and Hong-Geller, E. (2014) Differential expression of small RNAs from *Burkholderia thailandensis* in response to varying environmental and stress conditions. *BMC Genomics* **15**: 385.
Sun, X., Zhulin, I. and Wartell, R.M. (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res.* **30**: 3662-3671.
Thomason, M.K., Fontaine, F., De Lay, N. and Storz, G. (2012) A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Mol. Microbiol.* **84**: 17-35.
Thorner, A.R., Walker, D.H. and Petri, W.A. (1998) Rocky Mountain spotted fever. *Clin. Infect. Dis.* **27**: 1353-1359.
Tjaden, B. (2008) TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res.* **36**: W109-113.
Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D.*, et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**: 950-956.
Trotochaud, A.E. and Wassarman, K.M. (2004) 6S RNA function enhances long-term cell survival. *J. Bacteriol.* **186**: 4978-4985.
Tu, N., (2015) Characterization of two novel gene regulatory systems in the zoonotic bacterium *Bartonella henselae*. In. Tampa, FL: University of South Florida, pp. 112.
Updegrove, T.B., Shabalina, S.A. and Storz, G. (2015) How do base-pairing small RNAs evolve? *FEMS Microbiol. Rev.* **39**: 379-391.
Van Houten, B. and Kad, N. (2014) Investigation of bacterial nucleotide excision repair using single-molecule techniques. *DNA Repair* **20**: 41-48.
Wagner, E.G. and Simons, R.W. (1994) Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* **48**: 713-742.
Wagner, E.G.H. and Romby, P., (2015) Small RNAs in Bacteria and Archaea: Who They Are, What They Do, and How They Do It. In: Adv. Genet. J.C.D. Theodore Friedmann and F.G. Stephen (eds). Academic Press, pp. 133-208.
Walker, D.H. (2003) Principles of the malicious use of infectious agents to create terror - Reasons for concern for organisms of the genus *Rickettsia*. *Ann Ny Acad Sci* **990**: 739-742.
Warrier, I., Hicks, L.D., Battisti, J.M., Raghavan, R. and Minnick, M.F. (2014) Identification of novel small RNAs and characterization of the 6S RNA of *Coxiella burnetii*. *PLoS One* **9**: e100147.

Washietl, S., Hofacker, I.L. and Stadler, P.F. (2005) Fast and reliable prediction of noncoding RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 2454-2459.
Wassarman, K.M. (2002) Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes. *Cell* **109**: 141-144.
Wassarman, K.M. (2007) 6S RNA: a small RNA regulator of transcription. *Curr. Opin. Microbiol.* **10**: 164-168.
Wassarman, K.M. and Storz, G. (2000) 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**: 613-623.
Waters, L.S. and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615-628.
Weinert, L.A., Welch, J.J. and Jiggins, F.M. (2009a) Conjugation genes are common throughout the genus *Rickettsia* and are transmitted horizontally. *Proc. R. Soc. B* **276**: 3619-3627.
Weinert, L.A., Werren, J.H., Aebi, A., Stone, G.N. and Jiggins, F.M. (2009b) Evolution and diversity of *Rickettsia* bacteria. *BMC Biol.* **7**: 6.
Weissenmayer, B.A., Prendergast, J.G., Lohan, A.J. and 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**: e17570.
Westermann, A.J., Gorski, S.A. and Vogel, J. (2012) Dual RNA-seq of pathogen and host. *Nat. Rev. Microbiol.* **10**: 618-630.
Whitworth, T., Popov, V.L., Yu, X.J., Walker, D.H. and Bouyer, D.H. (2005) Expression of the *Rickettsia prowazekii* pld or tlyC gene in *Salmonella enterica* serovar Typhimurium mediates phagosomal escape. *Infect. Immun.* **73**: 6668-6673.
Winkler, H.H. (1976) Rickettsial permeability. An ADP-ATP transport system. *J. Biol. Chem.* **251**: 389-396.
Woodard, A. and Wood, D.O. (2011) Analysis of convergent gene transcripts in the obligate intracellular bacterium *Rickettsia prowazekii*. *PLoS One* **6**: e16537.
Woodward, T.E. (1973) A historical account of the rickettsial diseases with a discussion of unsolved problems. *J. Infect. Dis.* **127**: 583-594.
Woolfit, M., Algama, M., Keith, J.M., McGraw, E.A. and Popovici, J. (2015) Discovery of putative small non-coding RNAs from the obligate intracellular bacterium *Wolbachia pipientis*. *PLoS One* **10**: e0118595.
Wright, P.R., Georg, J., Mann, M., Sorescu, D.A., Richter, A.S., Lott, S.*, et al.* (2014) CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res.* **42**: W119-123.
Wright, P.R., Richter, A.S., Papenfort, K., Mann, M., Vogel, J., Hess, W.R.*, et al.* (2013) Comparative genomics boosts target prediction for bacterial small RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **110**: E3487-3496.
Xiao, B., Li, W., Guo, G., Li, B., Liu, Z., Jia, K.*, et al.* (2009) Identification of small noncoding RNAs in *Helicobacter pylori* by a bioinformatics-based approach. *Curr. Microbiol.* **58**: 258-263.
Zhang, J.Z., Hao, J.F., Walker, D.H. and Yu, X.J. (2006) A mutation inactivating the methyltransferase gene in avirulent Madrid E strain of *Rickettsia prowazekii* reverted to wild type in the virulent revertant strain Evir. *Vaccine* **24**: 2317-2323.

Vita

Mr. Casey Schroeder was born during 1983 in Iowa. He obtained a Bachelor of Science in clinical laboratory medicine from the University of Iowa and a Master of Science in biology from the University of Nebraska at Kearney. After graduating, Mr. Schroeder worked as a medical laboratory scientist at Mercy Medical Center in Des Moines, Iowa, earning specialist certifications in microbiology and laboratory safety from the American Society for Clinical Pathology (ASCP). Mr. Schroeder also served as a medical service corps officer in the Iowa Army National Guard.

Education
2005: B.S., Microbiology, University of Iowa, Iowa City, IA
2008: B.S., Clinical Laboratory Sciences, University of Iowa, Iowa City, IA
2011: M.S., Biology, University of Nebraska at Kearney, Kearney, NE
2016: PhD, Pathology, University of Texas Medical Branch, Galveston, TX

Board Certifications
2008: Medical Laboratory Scientist (ASCP)^{CM} #227146
2011: Specialist in Microbiology (ASCP)^{CM} #3054
2011: Specialist in Laboratory Safety (ASCP)^{CM} #222