Molecular analysis of the transfer (*tra*) genes of *Rickettsia bellii* RML 369-C and their function in bacterial conjugation

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Chan Chin Heu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
[MASTER OF (ARTS OR SCIENCE)] [DOCTOR OF (PHILOSOPHY/EDUCATION)]

Timothy J. Kurtti, Ulrike G. Munderloh

February 2016

Acknowledgements

I would like to acknowledge all the support that I received while working to achieve my dissertation. I would like to begin by thanking my parents, especially wonderful and nurturing mother, Yeng Xiong, who continued to provide love, advice, an ear to listen, and a nurturing home for me to return to if need be. I would also like to acknowledge my advisors, Drs. Timothy J. Kurtti and Ulrike G. Munderloh, for all their support and guidance, for which without them, it would be impossible to achieve my Ph.D. degree in entomology and gain the skills that I have today, and for being great role models. Furthermore, I would like to acknowledge my committee members, Drs. Ann M. Fallon and Peter J. Southern, my family (my sisters: Merci, Nancy, Tracy, Cindy, and Pa Heu; my brothers: Fong, Vayilen, and Vivson Heu; my brother-in-laws: Jonathan Ackerman, Chao Lee, and Dennis Thao; my cousins: Kue and Jarry Her, my extended family in Florida and here in Minnesota, and my uncles: Hugh and Bee Heu) for all the endless support and effort to make me feel at home in Minnesota and many more; my laboratory mates (Jonathan D. and Hyosook Oliver, Geoffrey E. Lynn, Adela Sarahi Olivia Chavez, Lisa D. Price, Nicole Y. Burkhardt, Roderick F. Felsheim, Curtis M. Nelson, and Michael J. Herron) for all the academic and professional support and advice; and my friends (Mai See Thao, Ka Ying Vang, Eric Yang, A Nancy Thao, Pay Thao, Chauru Huang, May Chang, and Woun Sian Souv and many others) for all their social support and encouragement. I would also like to thank everyone else who had supported me in one way or another, or at one time or many others. I would never have reached this point in my life without all of you. Thank you.

Dedication

I would like to dedicate this dissertation to my loving and nurturing mother, Yeng Xiong, who never gave up on raising nine children by herself and to the best of her ability.

Abstract

Rickettsia bellii is an obligate intracellular bacterium. Among the *Rickettsiales*, *R. bellii* is unusual because it encodes a complete set of transfer (*tra*) genes and pili-like structures associated with bacterial conjugation. To investigate the relationship of the *tra* genes with bacterial conjugation, I first characterized the *tra* genes into multiple operons and predicted the cellular localization of each of the Tra proteins using predicted transmembrane-spanning domains and signal peptides in the context of current knowledge of the bacterial conjugation system used by *E. coli*. I also characterized the transcriptional dynamics of *R. bellii tra* genes in comparison to stably transcribed genes to understand when the *tra* genes are actively transcribed during the rickettsial life cycle. I determined that the best reference genes, out of 10 tested, were methionyl tRNA ligase (*metG*) or a combination of *metG* and ribonucleoside diphosphate reductase 2 subunit beta (*nrdF*), using statistical algorithms from two different programs: Normfinder and BestKeeper. The transcription of *tra* genes was positively correlated with one another and up-regulated from 12 to 72 hours post inoculation (HPI) when compared to RBE_0422 (an inactivated transposase-derivative found within the *tra* cluster) suggesting that bacterial conjugation may occur at later exponential growth or early stationary growth. Furthermore, a complementation assay was designed using $traD_F$ of *R. bellii* to rescue a *traD* mutant *E. coli* in transferring DNA. However, $TraD_F$ expression was not detected in two different *E. coli* expression strains. The data suggest that future experiments to express rickettsial Tra proteins in *E. coli* should be undertaken to exploit the *R. bellii* bacterial conjugation system and connect the rickettsial Tra proteins into functional

genetic transfer systems.

Table of Contents

List of Tables

List of Figures

Chapter 1: Molecular characterization of the *tra* genes in the genus *Rickettsia* and their role in lateral gene transfer

Introduction

Rickettsiae are diplococcoid to rod-shaped, obligate intracellular alpha-proteobacteria. While many of which are avirulent (nonpathogenic) and symbiotically associated with arthropods, rickettsiae are mostly known for the diseases that they cause. The first rickettsia to cause a public outcry was reported in the 1870s where it caused "black measles", later known as Rocky Mountain Spotted Fever (RMSF), among human residence of the Bitterroot Valley of Montana. As this area began to industrialize with trading posts, apple orchard farming, gold mining, and lumber production, the cases of "black measles" increased with a fatality rate of ~70%, enough to terrorize residents and threaten economic progress in Montana. Howard Taylor Ricketts, a pathologist from University of Chicago, was recruited to research on the transmission of "black measles" in an attempt to identify control measures. Ricketts demonstrated that the causative agent was a diplococcoid bacterium transmitted by ticks to mammals and/or humans via a tick bite, demonstrating for the first time that an arthropod can transmit a bacterial pathogen. This revolutionized the field of vector-borne diseases (Harden 1990). Among the animal models Ricketts used, guinea pigs infected with *R. rickettsii* displayed measurable symptoms that included high fever and swelling of the scrotum, making them a good animal model for studying rickettsial infection. Ricketts fed wild caught ticks on guinea pigs and determined that only a small percentage of the tick population carried virulent *R. rickettsii*. Years after these discoveries, the bacteria were characterized and named *Dermacentroxenus rickettsi* after Dr. Ricketts and subsequently renamed *Rickettsia*

rickettsii (Harden 1990). When first proposed, the concept of a tick-borne bacterial pathogen was not widely accepted in Montana, since most tick bites did not result in disease and the east and west side of the valley showed a stark difference in spotted-fever cases. Nevertheless, these discoveries highlighted the medical importance of rickettsiae promoting further investigations in rickettsiology.

Medical Importance of the genus *Rickettsia*

The impact of virulent rickettsiae on public health in the U.S. extends beyond the Bitterroot Valley of Montana, and there are several other important pathogenic rickettsiae. Historically, rickettsiae were divided into spotted fever and typhus groups (SFG and TG, respectively) based on intracellular motility, presence or absence of rickettsial outer membrane protein A (OmpA) and the arthropod vector. *Rickettsia bellii* was identified to represent the ancestral group (AG) based on 16s rRNA and 23s rRNA (Stothard, Clark and Fuerst 1994). Furthermore, fifteen genes from *R. felis* chromosome and their homologues in nine other species of *Rickettsia* and two species of *Wolbachia* as outgroup identified the transitional group (TRG) in rickettsial phylogeny (Gillespie *et al.* 2007). Recently, Gillespie and colleagues proposed a more detailed typing scheme using whole genome analysis to identify core genes of *Rickettsia* species (Gillespie *et al.* 2008). Among the SFG rickettsiae, *Rickettsia parkeri* (Paddock *et al.* 2004) and *R. philipii* 364D (Shapiro *et al.* 2010; Parola *et al.* 2013) have recently been recognized to cause symptoms like maculopapular or petechial rash and fever similar to RMSF, which is caused by *R. rickettsii*. However, *R. parkeri* and *R. philipii* 364D infections are commonly associated with an eschar at the site of the tick bite and this is distinct from the general observations in RMSF. Because of these clinical symptom similarities, reported

RMSF cases may have erroneously included other rickettsioses (Raoult and Paddock 2005) overshadowing the importance of other rickettsiae. Rickettsial pox is caused by *R. akari* transmitted by mites and was first reported in Queens, NY, but later recognized as a global disease (Radulovic *et al.* 1996; Choi *et al.* 2005; Paddock *et al.* 2006; Zavala-Castro *et al.* 2009). *Rickettsia felis*, (a member of the TRG) and *R. typhi* (TG) are transmitted by fleas, *Ctenocephalides felis* and *Xenopsylla cheopis*, respectively, and contribute to rickettsiosis in the U.S., particularly in California, Hawaii, and Texas where molecular and immunological detection of both rickettsiae were found in the same species of arthropod and in the same geographical area (Williams *et al.* 1992; Schriefer *et al.* 1994a, 1994b; Manea *et al.* 2001; Civen and Ngo 2008; Adjemian *et al.* 2010). Rickettsiosis caused by *R. felis* and *R. typhi* extends across the globe with murine typhus occurring in China, Australia, Thailand, Kuwait, Tunisia and the Mediterranean area (Angelakis *et al.*; Azad 1990) and *R. felis* rickettsiosis occurring in all continents except Antarctica (Zavala-Velázquez *et al.* 2000; Znazen *et al.* 2006; Sunyakumthorn *et al.* 2008; Reif and Macaluso 2009; Renvoisé, Joliot and Raoult 2009; Parola 2011). Both diseases include mild febrile illness with rash and myalgia (Azad 1990; Parola 2011). Epidemic typhus, in contrast, is caused by *Rickettsia prowazekii* which is transmitted by human body lice and shares similar clinical signs with *R. felis* and *R. typhus* rickettsiosis, but differs in the severity of the disease (Bechah *et al.* 2008). It is a disease of historical impact, decimating and incapacitating entire armies during World War I and plaguing concentration camps during World War II (Bechah *et al.* 2008; Patterson 2012). The disease usually breaks out in areas of unsanitary conditions and dense population where frequent human contact occurs (Bechah *et al.* 2008). There is also a chance of

recrudescent infection by *R. prowazekii* known as Brill-Zinsser disease (Murray *et al.* 1950). A few other rickettsioses that have been recognized as medically important include Mediterranean spotted fever (*Rickettsia conorii*), African tick-bite Fever (*Rickettsia africae*), and Japanese spotted fever (*Rickettsia japonica*) where each disease is associated with the geographical location of the cases (Dumler 2012). Thus, rickettsiae have been a scourge of mankind through history by causing diseases in different parts of the world.

A general sylvatic and urban (domestic) cycle of *Rickettsia* and insights into the route of rickettsial infection in ticks

Human rickettsiosis is the result of transmission of rickettsiae via a hematophagous arthropod feeding on accidental hosts, i.e., humans. Transmission cycles for each rickettsial species are different because of the different vertebrate and arthropod hosts involved.

Among arthropod vectors, ticks and lice are of greatest importance, and in this thesis I focus on ticks as the arthropod vector. Hard ticks have four life stages: egg, larva, nymph, and adult (Troughton and Levin 2007). Larvae hatch from eggs, take a blood meal and molt into nymphs that take a second blood meal before molting to adults which then take a third blood meal and copulate before oviposition. Each molt and oviposition requires a blood meal from vertebrate hosts and ticks can acquire pathogenic rickettsiae from feeding on rickettsemic hosts. Larvae play a crucial role in the sylvatic cycle of *R. rickettsii* when they first acquire rickettsiae from feeding on an infected host as the infection persists essentially throughout the other life stages, and is passed transovarially to the next generation. The infected nymphs and adults will then transmit pathogenic

rickettsiae to vertebrates horizontally when feeding and possibly infect naïve ticks cofeeding on the same vertebrate host as has been seen with tick borne encephalitis virus (Randolph 2011). Although humans become infected with pathogenic rickettsiae, they are dead-end hosts and do not contribute to the distribution and maintenance of the rickettsiae in nature. In contrast, reservoirs for *R. rickettsii* such as ground squirrels or mice support asymptomatic infections and directly contribute to the distribution of rickettsiae.

In ticks, rickettsiae persist throughout the life stages via transstadial transmission, between molts, and by transovarial transmission, from mother to offspring. Nevertheless, *R. rickettsii* causes high mortality in ticks and only a small fraction of the infected offspring will survive (Niebylski, Peacock and Schwan 1999). These lethal effects on the vector encourage the horizontal transmission and infection of reservoir hosts, while transovarial transmission will likely select for an avirulent strain (Niebylski, Peacock and Schwan 1999). *Rickettsia prowazekii,* the agent of epidemic typhus, has also been shown to reduce the survival rate of infected lice (Snyder and Wheeler 1945), and thus to reduce the vectorial capacity of its arthropod host. In contrast, *Rickettsia bellii*-like rickettsiae that infect white flies have been shown to accelerate pre-adult development rate, as well as to decrease susceptibility of white flies to pathogens of hemiptera (Hendry, Hunter and Baltrus 2014), e.g. *Pseudomonas syringae*. Depending on the rickettsial species and whether they are pathogens or symbionts, the relationship with their host can be mutualistic (*Rickettsia peacockii* in *D. andersoni* (Niebylski *et al.* 1997b)), commensal, or parasitic (*R. rickettsii* and *R. prowazekii*). The relationship between host and symbiotic

(avirulent) rickettsiae is determined by multiple variables and many mechanistic details remain to be resolved.

A general model for the natural infection of ticks by pathogenic rickettsiae involves uptake of rickettsemic blood, infection of the midgut, and then spread to other organs. Originally, it was thought that rickettsiae invaded, propagated in, and then escaped from epithelial cells of the midgut to infect hemocytes in the hemolymph. Infected hemocytes would then facilitate spread of rickettsiae to infect the rest of the tick organs, i.e. salivary glands, fat body, ovaries, and tracheae (Socolovschi *et al.* 2009). However, tracking of fluorescent protein expressing rickettsiae in live tick tissues revealed that they used cells lining tracheae and tracheoles as conduits of dispersal within the tick (Baldridge *et al.* 2007b). Infection of the salivary glands facilitates infection of vertebrate hosts, and infection of the ovaries may facilitate transovarial transmission. Both infection routes perpetuate the natural transmission of rickettsiae. Identification of key mechanisms of rickettsial dissemination within ticks requires additional experimental study.

Phylogeny and taxonomy of the genus *Rickettsia*

Historically, the *Rickettsiales* was taxonomically delineated based on biological, chemical, and morphological data. Subsequently, molecular data have been used to revise the taxonomy of the *Rickettsiales* using the sequence of the *16s rRNA* before the advent of whole genome sequencing. However, the growing number of whole genome sequences of rickettsiae in the genomic era solidified the revision of the taxonomic relationships within the *Rickettsiales*. It still stands that all members of *Rickettsiales* are obligate intracellular bacteria and includes the families, *Holosporaceae*,

Anaplasmataceae and *Rickettsiaceae* (Emelyanov 2001). The family *Rickettsiaceae* is further divided into the genera *Orientia* and *Rickettsia*. *Orientia* was once grouped together with *Rickettsia*, but differences in genetics and metabolic pathways have subsequently distinguished *Orientia* from *Rickettsia* (Tamura *et al.* 1995). The genus *Rickettsia* is split into four subgroups based on whole genome phylogenetic analysis (Gillespie *et al.* 2010): the ancestral group (*R. bellii* RML 369-C (Ogata *et al.* 2006) and *Rickettsia canadensis* McKiel (Eremeeva *et al.* 2005)); the transitional group (*Rickettsia felis* URRWXCal2 (Ogata *et al.* 2005a, 2005b), *Rickettsia akari* Hartford, and *Rickettsia australis* Cutlack); the spotted fever group (*Rickettsia rickettsii* Iowa (Ellison *et al.* 2008), *Rickettsia conorii* Malish (Ogata *et al.* 2001), *Rickttsia peacockii* Rustic (Felsheim, Kurtti and Munderloh 2009), *Rickttsia massiliae* MTU5 (Blanc *et al.* 2007a), *Rickettsia buchneri* ISO-7 (Gillespie *et al.* 2012; Kurtti *et al.* 2015)), and the typhus group (*Rickettsia typhi* Wilmington (Mcleod *et al.* 2004) and *Rickettsia prowazekii* Madrid E (Andersson *et al.* 1998)).

The study of nonpathogenic (avirulent) rickettsiae facilitates greater comprehension of rickettsial pathogenicity mechanisms

Because of both human and animal health concerns linked to pathogenic rickettsiae and the potential use of rickettsiae for biological warfare, some studies have been conducted with pathogenic rickettsiae but study of nonpathogenic rickettsiae has been largely neglected. Avirulent rickettsiae can generate important information about their pathogenic counterparts. For example, genomic comparisons have consistently shown that the majority of genes are shared by close relatives, such as *R. rickettsii* Sheila Smith and the avirulent *R. rickettsii* Iowa and *R. peacockii* Rustic. However, several

genes found in pathogenic rickettsiae are inactivated or missing from nonpathogenic rickettsiae indicating that these are potential virulence genes (Ellison *et al.* 2008; Felsheim, Kurtti and Munderloh 2009). Furthermore, comparing the genomes of pathogens to other nonpathogenic bacteria from other genera can illuminate the evolution of pathogenicity. Analysis of the genome of *R. prowazekii* Madrid E first revealed an evolutionary link with mitochrondria, with indications of a shared pathway from free-living to obligate intracellular bacteria, and reductive genome evolution (Andersson *et al.* 1998). The hypothesis of reductive genome evolution suggests that many of the genes important for basic functions of the organism are lost and a dependence on host metabolites becomes necessary to retain viability. Possibly, the loss of genes in important metabolic pathways requires that pathogenic bacteria parasitize host metabolites at a greater rate causing more harm to the host. Therefore, using nonpathogenic rickettsiae to investigate the genetics of pathogenic rickettsia can be a practical approach to identifying genes that are of great interest and value, perhaps pointing to a pivotal role in the biology and evolution of rickettsiae.

In addition, the genomic attributes of pathogenic and nonpathogenic rickettsiae show stark contrasts worth investigating. The genomes of nonpathogenic rickettsiae tend to be larger than those of pathogenic ones (Gillespie *et al.* 2012). The different genome sizes are attributed to mobile genetic elements (MGE) acquired through lateral gene transfer (LGT) that may still be functional in further dissemination (Ogata *et al.* 2005a, 2005b, 2006; Blanc *et al.* 2007a; Gillespie *et al.* 2014, 2015). Insertion sequences (IS) are a part of the MGE that contribute to amplification and replication of IS, genome shuffling (Felsheim, Kurtti and Munderloh 2009), and insertion of new genetic material (Gillespie

et al. 2012). This could lead to acquisitions of new genes that promote a mutualistic or parasitic relationship with the host depending on the genes acquired. LGT apparently has had a profound impact on the genomes of rickettsiae and potentially determines pathogenic potential and the basis for host-symbiont interactions. .

Lateral gene transfer in the genus *Rickettsia*

The mechanisms of LGT are natural transformation, transduction, and/or bacterial conjugation. Natural transformation occurs when a bacterium acquires genetic material from the environment and naturally incorporates the genetic material into its genome. In contrast, when a defective virus carries and injects bacterial DNA into a bacterium leading to the incorporation of new DNA into the bacterial genome it is called transduction. The third LGT event is bacterial conjugation, where a bacterium (donor) transfers genetic material via direct contact with another bacterium (recipient). All three mechanisms may be facilitated in rickettsiae.

The obligate intracellular lifestyle of rickettsia complicates these LGT events because the host cell membrane acts as a barrier to the introduction of new DNA. However, careful observation of the genomes of rickettsia reveals evidence of LGT events. The genome sequence of *R. bellii* RML 369-C (Ogata *et al.* 2006) contains evidence of LGT. The *R. bellii* genome contains genes that are found in the gamma-proteobacteria *Legionellaceae* and *Parachlamydiaceae.* This finding has generated the hypothesis that their coinfection of an amoeba "melting pot" promoted LGT events between intracellular microorganisms capable of eluding digestion by amoeba (Moliner, Fournier and Raoult 2010). Active ingestion of bacteria by amoebae gathers bacteria into this intracellular niche enabling physical contact between bacteria that escaped digestion

and can lead to bacterial conjugation. It is also possible that the DNA from partially digested bacteria can be acquired, leading to natural transformation. The amoeba model represents a single eukaryotic cell arena that supports LGT between bacteria.

Ticks may also be a “melting pot” by facilitating the interaction between environmentally acquired bacteria and the natural tick microbiome. The microbiome of ticks can change in composition, abundance, and diversity according to abiotic and biotic factors. For example, the composition of the microbiome changes with the seasons in *Rhipicephalus turanicus* and *R. sanguineus* (Lalzar *et al.* 2012). The different life stages and nutritional states of *Ixodes scapularis* also demonstrated significant differences in their microbial composition (Moreno *et al.* 2006; Narasimhan *et al.* 2014). Further differences include variations among the the bacterial genera in different species of ticks. However, ticks may share some bacterial genera based on common life style such as diet, i.e. blood, or host. In fact, *Pseudomonas* and *Serratia* were found in most studies of tick microbiomes, suggesting either contamination and/or true bacterial colonization of ticks by these two ubiquitous bacterial genera (Moreno *et al.* 2006; Clay *et al.* 2008; Andreotti *et al.* 2011; Lalzar *et al.* 2012; Narasimhan *et al.* 2014; Narasimhan and Fikrig 2015). Interestingly, species of *Serratia* and *Pseudomonas* have been identified in mosquitoes and the presence of *Serratia odorifera* in the midgut enhances the susceptibility of *Aedes aegypti* to dengue virus (Apte-Deshpande *et al.* 2012). The microbes present in specific tick tissues may help to identify key patterns that identify important LGT events. For instance, three rickettsial species were identified in the midgut of *Amblyomma maculatum*, but only *R. parkeri* was identified in the salivary glands and the detection of *R. parkeri* in the saliva indicates transmission of *R. parkeri* via tick bite (Budachetri *et al.*

2014). Further, *Francisella* was identified in the midgut and salivary gland of *A. maculatum*, but no detectable DNA of *Francisella* was found in the saliva suggesting coinfection of the same organs occupied by *Rickettsia* spp., but no apparent horizontal transmission (Budachetri *et al.* 2014). The coinfection of organs, i.e. midgut and ovaries, may provide a "melting pot" environment that facilitates physical contact of different bacterial genera possibly leading to LGT.

LGT events in *Rickettsiales* have occurred throughout the evolution from free-living to obligate intracellular life. LGT events may have occurred in the ancestral state with multiple rates of degradation of the acquired genes by different rickettsiae (Fuxelius *et al.* 2008). Gene gain events have been calculated to have occurred before speciation of the spotted-fever group rickettsia. Two gene gain events in the transitional group after speciation, and the high frequency of gene gain in *R. bellii* resulted in the prediction that its genome was larger than that of its ancestor (Georgiades *et al.* 2011). In fact, the selfish insertion sequence, ISRpe1, has been documented to invade different genera of bacteria including *Rickettsia* species throughout evolutionary time via LGT (Duron 2013). The invasion of ISRpe1 in *R. peacockii* was responsible for the genome shuffling of *R. peacockii* demonstrating that LGT may have an impact on the speciation of rickettsiae (Fuxelius *et al.* 2008; Felsheim, Kurtti and Munderloh 2009; Georgiades *et al.* 2011). Further investigation of LGT in rickettsiae revealed that the genome of *R. felis* displayed high incongruity for its genome implying that *R. felis* is a chimeric species that has acquired most of its genes from various sources (Merhej *et al.* 2011). Thus LGT events have shaped the evolutionary history of rickettsiae.

The genome sequences of *Rickettsia* spp. continue to display evidence of LGT that may be beneficial to the rickettsiae (Ogata *et al.* 2006; Blanc *et al.* 2007a; Felsheim, Kurtti and Munderloh 2009; Gillespie *et al.* 2012). For example, *Rickettsia peacockii* maintains five genes on the pRPR plasmid that are distantly related to the *Pseudomonas aeruginosa* glycosylation island and may function in the uptake of dihydroxyacetone phosphate for phospholipid biosynthesis (Felsheim, Kurtti and Munderloh 2009). Recently, an RTX type I secretion system (TISS) was found in *R. felis* plasmid pLSU-Lb that was closely related to the TISS genes of a *Cardinium* endosymbiont cBtQ1 of *Bemisia tabaci*, in which the RTX TISS genes were thought to originate through lateral gene transfer from *Vibrio* species. In addition, *R. buchneri* encodes a complete biotin operon on a plasmid (pREIS2) that shares homology with biotin operons recently acquired by horizontal gene transfer in *Neorickettsia risticii*, *Neorickettsia sennetsu* and *Lawsonia intracellularis*. This operon also shares homology with the biotin operon discovered in *Wolbachia w*Cle that was shown to synthesize and supplement necessary vitamin B7 to bed bugs (Nikoh *et al.* 2014). The presence of high colinearity of genes and group of genes involved in LGT suggests that the genes acquired can still be functional.

Tra genes

The presence of transfer (*tra*) genes in some rickettsial genomes is considered major evidence for LGT. Bioinformatics predict that the *tra* genes in rickettsiae function in bacterial conjugation. The general structure of the conjugation machinery in bacteria consists of the pilus, the core complex, the inner membrane platform, and the cytoplasmic ATPases that constitute the mating pair formation apparatus. An *E. coli* model of pilus biogenesis indicates that propilin units are recruited to the inner membrane where the

leader sequence is cleaved leaving the mature pilin integrated in the inner membrane (Cabezon *et al.* 2015). Cytoplasmic ATPases hydrolyze ATP for energy to form the pilus by incorporating pilins helically into a tube. The pilus then attaches to a recipient bacterium and retracts to bring the recipient bacterium to the donor. During this process, the core complex changes conformation during ATP hydrolysis and forms a bridge between the inner and outer membrane with proteins involved in conjugation. The transfer DNA, processed by the relaxasome and with the help of other transfer proteins, finds and docks the mating pair formation complex and transfers the DNA strand to the recipient. Bacterial conjugation is concluded with the separation of the cells and the replication and maintenance of the DNA strand in the transconjugant. The *tra* genes of *R. bellii* have been predicted to function in the essential aspects of bacterial conjugation (Ogata *et al.* 2006).

The *tra* genes of *R. bellii* consist of two parts: in one part the majority of the *tra* genes are orthologous to the F type IV secretion system (T4SS) of γ-proteobacteria, and in the other part genes are orthologous to the tumor inducing (Ti) plasmid of *Agrobacterium tumefaciens* (Ogata *et al.* 2006). The F-like region consists of *traL*, *-E*, *-B*, *-V*, *-C*, *-W*, *-U*, *-N*, *-F*, *-H*, *-G*, *-D*, and *trbC*, all of which, except *traD*, contribute to the mating pair formation during bacterial conjugation (Fig. 1.1). In *E. coli*, TraL, -E, -B, -V, -C, -W, -U, -F, -H, -G, and TrbC function in pilus biogenesis because mutations in each gene affect the pili and inhibit or reduce infection by bacteriophages (Willetts 1973; Frost, Paranchych and Willetts 1984; Frost, Ippen-Ihler and Skurray 1994; Firth, Ippen-Ihler and Skurray 1996). TraL, -E, -B, -C, and –G are found integrated in the inner membrane, while TraW, -U, -F, and TrbC have been experimentally verified to localize

to the periplasm (Fig. 1.1) (Frost, Ippen-Ihler and Skurray 1994; Firth, Ippen-Ihler and Skurray 1996). TraG and TraN function in mating pair stabilization with TraG also functioning in entry exclusion of the transfer DNA (Manning, Morelli and Achtman 1981; Firth and Skurray 1992; Maneewannakul, Kathir and Ippen-Ihler 1992; Klimke *et al.* 2005; Marrero and Waldor 2005, 2007; Audette *et al.* 2007). TraD is a coupling protein and is found localized to the inner membrane (Panicker and Minkley Jr 1992) where it assists in DNA processing and transfer by the relaxosome in *E. coli* (Lang *et al.* 2011). Missing or yet to be identified in *Rickettsia* species are *traJ* (involved with transcription regulation), *traM* and *traY* (involved in DNA metabolism and transcription regulation), *traX* (pilin acetylation role), *traQ* (pilin localization), *traS* and *–T* (surface exclusion). The Ti-like region consists of $traA_{Ti}$ and $traD_{Ti}$ that constitutes the relaxosome where $traA_{Ti}$ encodes the relaxase (similar in function to TraI in *E. coli*) while $TraD_{Ti}$ assists the relaxase in processing the DNA for genetic transfer (Cho and Winans 2007). However, the $traD_{Ti}$ gene is truncated in *Rickettsia* species and may not be translated to a functional, full-length $TraD_{Ti}$. By comparison, $TraD_F$ from the F-like region is predicted to translate a 562 amino acid protein with functional domains that may interact with $TraA_{Ti}$. The *tra* genes may originate from different bacterial species, but altogether the *tra* genes in *Rickettsia* fulfill the requirements for essential components of bacterial conjugation - the mating pair formation and the DNA-processing relaxosome.

Taken together, the bioinformatics data indicate that rickettsial *tra* genes can mobilize genetic elements including the *tra* genes to other rickettsiae. A screening of the *tra* genes in rickettsiae isolates shows that at least one of the *tra* genes has been found in several rickettsia members, suggesting that *tra* genes are common within this group

(Weinert, Welch and Jiggins 2009). *Rickettsia bellii* RML 369-C and *R. massiliae* MTU5 have the complete *tra* cluster (Ogata *et al.* 2006; Blanc *et al.* 2007a) and the more recently sequenced genomes *of Rickettsia buchneri,* also known as Rickettsial endosymbiont of *Ixodes scapularis* (REIS), identified more variations of *tra* clusters, highlighting the difference between the *R. bellii* and *R. masilliae tra* clusters (Gillespie *et al.* 2012). *Rickettsia buchneri* also shows evidence of the *tra* cluster in a native plasmid (pREIS3) that shares homology with *R. felis* LSU-Lb plasmid (pLbaR) supporting the mobility of the *tra* genes. This evidence supports a more recent acquisition of the *tra* genes. Thus, the *tra* genes provide genetic and predicted functional evidence of LGT events.

Despite this evidence, the requirement of physical contact between the donor and recipient bacteria needed for bacterial conjugation excludes other extracellular bacteria for the obligate intracellular rickettsia. As previously mentioned, the amoeba "melting pot" hypothesis is one explanation of how an ancestor of *Rickettsia* acquired the *tra* genes. If LGT had occurred more recently between *Rickettsia* spp., then the physical barrier must have been broken down. Rickettsial coinfection of the same host (Carmichael and Fuerst 2010) where physical contact can occur in different tissues, e.g. hemolymph, salivary glands, or tracheae, could be a potential pathway for LGT. The detection of different genera of bacteria infecting the same host (Eremeeva *et al.* 2009; Milhano *et al.* 2010; Lommano *et al.* 2012; Schicht, Schnieder and Strube 2012) argues that coinfection occurs more often if screened for it. This, however, does not rule out the possibility of bacterial conjugation occurring in the same host cell. For example, the horizontal transfer of plasmid pNH4 from *Bartonella rattaustraliani* AUST/NH4 to

Rhizobium radiobacter was demonstrated using the type IV secretion system (T4SS) within an amoeba (Saisongkorh *et al.* 2010). However, the transconjugant *R. radiobacter* could not maintain the plasmid stably and the plasmid was lost after several subcultures. This experiment demonstrates that if two bacteria coinfect the same cell with one containing the appropriate transfer machinery, they can exchange genetic material. Thus, indicating that LGT via bacterial conjugation in *Rickettsia* spp. can occur both within the cell or host.

Plasmids

When LGT occurs, the genetic material that usually transfers is a plasmid, circularized DNA that is usually smaller in size than the main chromosome. The first plasmid ever detected within *Rickettsia* species is the plasmid pRF, revealed via whole genome sequencing of *R. felis* (Ogata *et al.* 2005a). Subsequently, transformation of *R. monacensis* using a transposon mechanism inserted the transposon in the native plasmid and revealed yet another rickettsial plasmid (pRM) (Baldridge *et al.* 2007a). Additionally, the genome sequence of *R. massiliae* revealed a plasmid (pRMA) and a complete *tra* gene cluster integrated in the chromosome (Blanc *et al.* 2007a). Moreover, Baldridge *et al.* (2010) discovered that plasmids are more common in *Rickettsia* than previously thought. The discovery of three native plasmids (pRAM18, pRAM23, and pRAM32) coexisting in *Rickettsia amblyommii* AaR/SC resulted in the basis for the creation of plasmid shuttle vectors for transforming *Rickettsia* species (Burkhardt *et al.* 2011). Many transformations using these shuttle vectors proved to be successful, as screened using fluorescent microscopy, implying that receiving exogenous rickettsial plasmids via LGT will allow the maintenance and proliferation of these plasmids as well. In addition,

Rickettsia buchneri have four plasmids total indicative of plasmid compatibility. Two of those plasmids (pREIS1 and 3) carry the complete cluster of *tra* genes (Gillespie *et al.* 2012) indicating that these plasmids may potentially be transferred to other rickettsiae naturally via bacterial conjugation. Interestingly, the F portion of the *tra* genes encoded by pLBaR, the plasmid of *R. felis* LSU-Lb (a rickettsial strain isolated from *Liposcelis bostrychophila* –the booklouse- that is neither hematophagous nor medically relevant) is perfectly colinear with the *tra* genes in the pREIS3 plasmid of *R. buchneri*, the rickettsial symbiont of the blacklegged tick (Gillespie *et al.* 2014). This implies that this *tra* gene cluster may have shared the same origin, possibly with a yet unidentified source or from either one of these rickettsiae through LGT using the rickettsial plasmid as a vehicle.

R. bellii as a model organism for the study of *tra* gene function

Rickettsia bellii is in the ancestral clade of the obligately intracellular family *Rickettsiaceae*. Electron microscopy reveals ultrastructural features shared by *Rickettsiaceae*; e.g. trilaminar cell wall and halo or slime layer (Philip *et al.* 1983), inner leaflet thicker than the outer leaflet of the outer membrane, and microcapsular layer (Labruna *et al.* 2004b). *Rickettsia bellii* has been isolated mainly from ticks (Table 1.1). Most reports of *R. bellii*-like organisms described molecular detection and sequence identity of partial fragments of several genes that mainly included the 16s rRNA gene, the *htrA* (17-kDa outer membrane protein) and *gltA* (citrate synthase) genes, and less often used genes, such as *atpB* (ATPase subunit alpha) and *coxA* (cytrochrome oxidase subunit I) (Table 1.2). Sequences were searched using BLAST (Basic Local Alignment Sequence Tool), and sequences that had high similarities with *R. bellii* isolates were considered *R. bellii*-like. Through sequence analysis of gene fragments amplified using

the polymerase chain reaction (PCR), *R. bellii* has been detected in a variety of host species making it the *Rickettsia* with the broadest host range to date when compared to other *Rickettsia* spp.

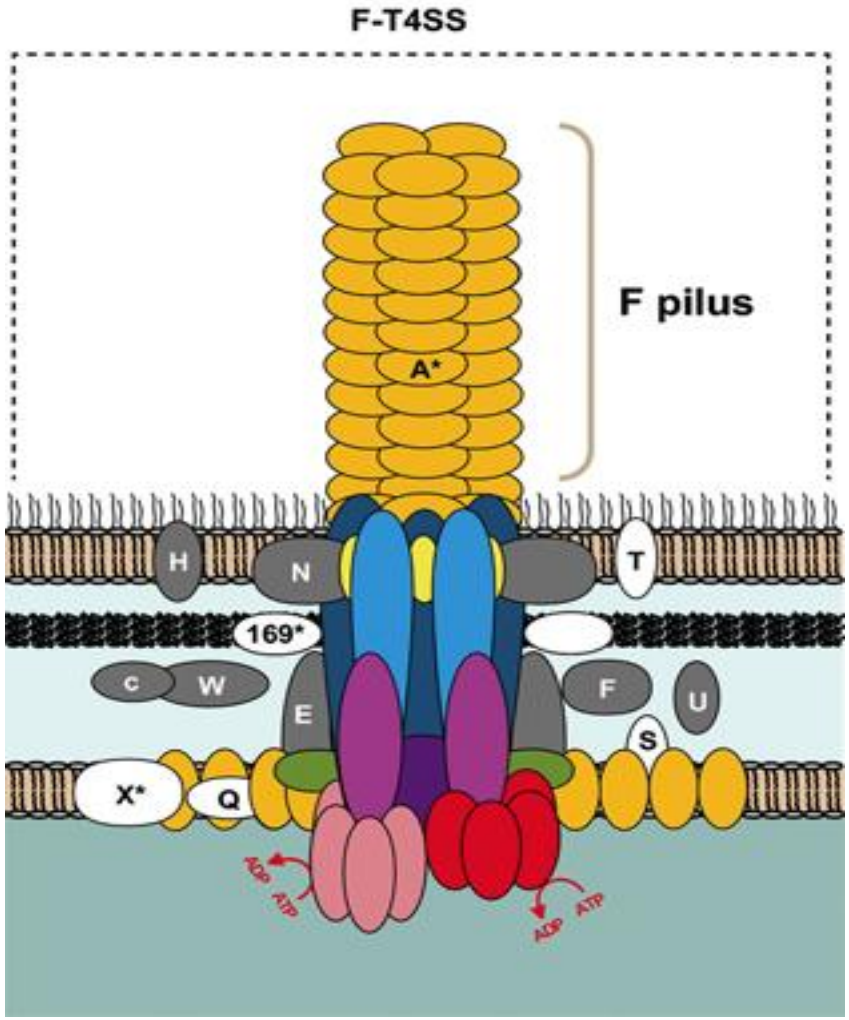
Rickettsia bellii RML 369-C is an ideal candidate to study rickettsial *tra* genes because it has a broad host range and encodes the complete set of *tra* genes (Ogata *et al.* 2006). Hosts include both hard and soft ticks (Table 1.1) (Philip and Casper 1981; Philip *et al.* 1983; Labruna *et al.* 2004a, 2004b), as well as insects (white flies (Gottlieb *et al.* 2006; Rao *et al.* 2012), hymenopteran parasitoids (Giorgini *et al.* 2010; Gebiola *et al.* 2012), and beetles (Lawson *et al.* 2001)). The large host range increases the chances of two *R. bellii* strains sharing the same host (Table 1.1 and 1.2), of rickettsiae making physical contact, and of subsequent LGT events. In fact, coinfections with *R. bellii* and another rickettsial species, e.g., *R. rhipicephali,* have been documented in *Dermacentor occidentalis* (Table 1.1), and *R. bellii* has also been shown to coinfect the same host together with a different bacterial genus, e.g. *Wolbachia* spp., in *Curculio sikkimensis* (Table 1.2) (Toju and Fukatsu 2011). In addition, *R. bellii* has been reported to superinfect a specimen of *Dermacentor variabilis* already infected with, *R. montanensis* and *R. rickettsii,* making *R. bellii* a candidate to examine potential bacterial conjugation between virulent and avirulent rickettsiae (Carmichael and Fuerst 2010). Although multiple strains of *R. bellii* carry a native plasmid that may be involved in gene transfer (Labruna *et al.* 2007a, 2007b; Baldridge *et al.* 2010), *R. bellii* RML 369-C does not carry this plasmid, which may have been lost during years of passages in cell culture (Baldridge *et al.* 2010). Nevertheless, the *tra* genes of *R. bellii* are a type of integrative, conjugative element that can transfer its own DNA element without the need for a

plasmid. The genome sequence of *R. bellii* RML 369-C has revealed a complete set of *tra* genes that are syntenic with *tra* genes in other *Rickettsia* spp. that carry a complete set, making this strain ideal for transcriptional and functional analyses of rickettsial *tra* genes.

In the following chapters, I characterize the *tra* genes of *R. bellii* RML 369-C. First, I have determined that the *tra* genes are transcriptionally active. In silico predictions of the organization of *tra* genes into operons, and predictions of the Tra protein domains using bioinformatics and knowledge of the F- type IV secretion system (F-T4SS) collectively translated into a structural model of the conjugation system of *R. bellii*. Second, relative quantification of the transcripts of the *tra* genes through time using qRT-PCR together with mathematical and statistical comparison to reference gene transcripts indicated the timing when rickettsial conjugation most likely would occur. Finally, complementation of mutant or deleted *tra* genes of the F-T4SS of *E. coli* with homologous *R. bellii tra* genes should identify their function, which will be critical for understanding the role of the *tra* genes in the biology of *Rickettsia* species.

Fig. 1.1. Structure of the F-T4SS of *R. bellii*. This graphical representation and legend of the F-T4SS of *R. bellii*, or represented by *R. buchneri Rickettsiales*-amplified genetic element (RAGE)-Be *tra/trb*, was taken and modified from Gillespie et al. 2014 and is not illustrated to scale. (A) A graphical display of the F-T4SS of *R. bellii* based on the *trb*/*tra* paradigm of *E. coli* plasmid F. (B) Comparison of gene synteny for F-T4SS (*R. bellii tra* cluster and *E. coli* plasmid F). Genes homologous across both systems are indicated by gray shading. Color schema for both (A) and (B) correspond with each other with the same color indicating the gene and its corresponding protein. Dashed box indicates the relaxase ($TraA_{Ti}$) and the conjugative $TraD_{Ti}$ protein of the Ti-like region. Black indicates non-T4SS gene.

A



B

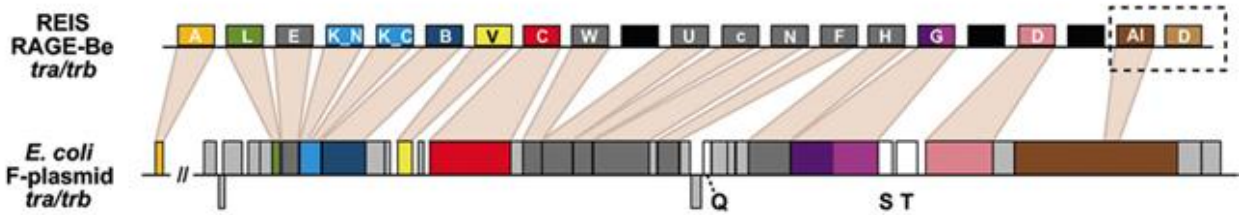

REIS
RAGE-Be
tra/trb
A
L
E
K_N
K_C
B
V
C
W
U
c
N
F
H
G
D
AI
D
E. coli
F-plasmid
tra/trb
//
Q
S T

Table 1.1. *Rickettsia bellii* and *R. bellii-like* organisms and their tick host(s)

Host(s)	***Rickettsia* isolate/ identification or detection method**	**Where the tick came from**	**Host shared by *Rickettsia* and/or others**	**Bacterial species sharing host**	**Co-infection**	**Source**
Dermacentor variabilis *D. andersoni, D. occidentalis, D. parumapertus, D. albopictus, Haemaphysalis leporispalustris, Argas cooleyi, Onithodoros concanensis*	*R. bellii* RML 369-C from *D. variabilis* Serology, cellular and animal pathogenicity, morphology, and protein profiling determination of *R. bellii*-like	Vegetation in Fayetteville, Arkansas Montana, North Carolina, Ohio, California, South Dakota, and Oklahoma	none	none	None	(Philip *et al.* 1983)
D. andersoni	*R. bellii*-like determined by PCR/restriction fragment length polymorphism	Vegetation in Bitterroot Valley, Montana and Poudre Canyon, Colorado	*D. andersoni*	*R. rickettsii, Francisella sp.*	potentially *Francisella sp.*	(Niebylski *et al.* 1997a)
Amblyomma cooperi	*R. bellii* Ac25 and 29 PCR confirmed using *htrA* and *gltA*	Farmlands in Pedreira Municipality, Sao Paulo	*A. cooperi*	R. sp. strain COOPERI	No	(Labruna *et al.* 2004b)
Amblyomma scalpturatum *Amblyomma humerale, A. ovale, A. rotundatum, A. oblongoguttatum,*	*R. bellii* As17, PCR confirmed using *gltA* PCR detection of *R. bellii*-like using *gltA*	Amazon forest in Rondonia, Brazil	none	none	No	(Labruna *et al.* 2004a)
D. variabilis	PCR detection of *R. bellii*-like using *htrA* and *gltA*	Ohio	*D. variabilis*	*R. rickettsii* and *R. montanensis*	Yes	(Carmichael and Fuerst 2010)
Ixodes loricatus	PCR detection of *R. bellii*-like using *gltA*	Opossum in Mogi das Cruzes, Sao Paulo State, Brazil	none	none	No	(Horta *et al.* 2006)

Amblyomma aureolatum	*R. bellii* Mogi PCR confirmed using *gltA*	Infested dogs in Taiacupeba, Mogi das Cruzes, Sao Paulo	*A. aureolatum*	*R. rickettsii* Taiacu	No	(Pinter and Labruna 2006)
Amblyomma neumanni	*R. bellii* AN4 PCR confirmed using *htrA* and *gltA*	Walking survey and dry ice trap in Establecimiento "La Luisiana", Dean Funes, Cordoba Province, Argentina	*A. neumanni*	Candidatus *R. amblyommii* AN13	No	(Labruna *et al.* 2007a)
Haemaphysalis juxtakochi	*R. bellii* HJ#1,2,3,4, and 7 PCR confirmed using *htrA* and *gltA*	Intervales State Park, Ribeirão Grande Municipality, state of São Paulo and from a brocket deer in Sao Paulo Municipality, SP	*H. juxtakochi*	*R. rhipicephali* HJ#5	No	(Labruna *et al.* 2007b)
Amblyomma ovale, A. incisum	Isolate of *R. bellii* from *A. ovale*, PCR detection of *htrA* and *gltA* in *A. incisum*	Vegetation in Intervales State Park, Sao Paulo	none	none	No	(Pacheco *et al.* 2008)
D. occidentalis	PCR detection of *R. bellii*-like using *gltA*	Southern California	*D. occidentalis*	*R. rhepicephali*	Yes	(Wikswo *et al.* 2008)
Amblyomma dubitatum	*R. bellii* Ad-CORD-SP-25 and Ad-INT, PCR confirmed using *htrA* and *gltA*	Vegetation and capybara in 16 states of Sao Paulo, Brazil	none	none	No	(Pacheco *et al.* 2009)
Amblyomma nodosum	*R. bellii* Pontal, PCR confirmed using *htrA* and *gltA*	Passeriform birds in Atlantic forest, eastern Brazil	*A. nodosum*	*R. parkeri* NOD	No	(Ogrzewalska *et al.* 2009)
Ixodes ricinus	PCR detection of *R. bellii* –like using 16s rRNA gene	Bijlmerweide, Duin en Kruidberg in Netherlands	*I. ricinus*	*R. conorii*, *R. helvetica*, R. SP. IRS, *R. typhus* group	No	(Sprong *et al.* 2009)
Amblyomma tigrinum	PCR detection of *R. bellii* –like using *htrA* and *gltA*	Moreno Department, Province of Santiago del Estero, Argentina	none	none	No	(Tomassone *et al.* 2010)

A. nodosum	*R. bellii* Pontal, PCR confirmed using *htrA* and *gltA*	Atlantic forest, eastern Brazil	*A. nodosum*	*R. parkeri* NOD	No	(Ogrzewalska *et al.* 2011)
I. ricinus and Dermacentor reticulatus	PCR detection of *R. bellii* -like using 16s rRNA gene	Ticks from flagging, deer, dog, and humans in England, Wales, and Scotland	*I. ricinus and D. reticulatus*	*R. helvetica*, *R. limoniae*-like, *R. typhi*-like, *R. massiliae*-like	yes, *R. massiliae*-like and *R. bellii*-like in *D. reticulatus*	(Tijsse-Klasen *et al.* 2011)
Amblyomma varium	PCR detection of *R. bellii*-like using *gltA*	Allpahuayo Mishana National Reserve near Iquitos, Peru	none	none	No	(Ogrzewalska *et al.* 2012)
Amblyomma sabanerae	PCR detection of *R. bellii*-like using *gltA*	El Salvador	none	none	No	(Barbieri, Romero and Labruna 2012)
Amblyomma cajennense	PCR detection of *R. bellii*-like using *gltA*	Indian population and vegetation in Comodoro, Brazil	*A. cajennense*	*R. amblyommii*	No	(de Barros Lopes *et al.* 2014)
Amblyomma longirostre	*R. bellii* Pontal, PCR confirmed using *htrA* and *gltA*	*Coendou insidiosus* (hairy dwarf porcupine) in the state of Bahia, northeast Brazil	*A. longirostre*	*R. amblyommii* str. Aranha	No	(McIntosh *et al.* 2015)
A. humerale, *Amblyomma naponense*	*R. bellii* Pontal, PCR confirmed using *htrA* and *gltA*	*A. humerale* on *Dasypus novemcinctus* and *Didelphis marsupialis*, *A. naponense* on *Tayassu pecari* in Mato Grosso and Para of Brazil within the Amazon Biome	*A. humerale*/ *A. naponense*	*R. amblyommii* and *R. felis*/ *R. africae*	No	(Soares *et al.* 2015)

Grayed organism names are of particular interest in this thesis. Genes used to confirm *R. bellii*-like agents were provided for future reference. The details of the geographical location and hosts were provided as accurately as described within the referenced articles.

Table 1.2. *Rickettsia bellii*-like organisms and their insect host

Host(s)	**Common name and/or Order:Family Detection of *R. bellii-like***	**Where the host was collected from**	**Host shared**	**Bacterial species sharing host**	**Co-infection**	**Source**
Brachys tessellatus	Buprestid leaf-mining beetle, Coleoptera: Buprestidae PCR detection of *R. bellii*-like using 16s rRNA	South Carolina	none	none	No	(Lawson *et al.* 2001)
Acyrthosiphon pisum str. *OY*	Pea aphid, Hemiptera: Aphididae PCR detection of *R. bellii*-like using 16s rRNA and *gltA*	Japan	*A. pisum*	*Buchnera*	yes	(Sakurai *et al.* 2005)
Tetranychus urticae	Two-spotted spider mite, Acari: Tetranychidae PCR detection of *R. bellii*-like using 16s rRNA gene	Prosser, WA; Grandview, WA; Lodi, CA	*T. urticae*	*Wolbachia*, *Caulobacter*	yes, *R. bellii*-like with *Wolbachia* or potentially with *Caulobacter*	(Hoy and Jeyaprakash 2005)
Bemisia tabaci	Sweet potato whitefly, Hemiptera: Aleyrodidae PCR detection of *R. bellii*-like using 16s rRNA gene	from B biotype lab colony originating from Israeli cotton fields	*B. tabaci*	"Ca. *Portiera aleyrodidarum*", "Ca. *Hamiltonella defensa*"	yes, all three	(Gottlieb *et al.* 2006)
Calvia quatuordecimguttata str. *I, Chrysopid species lacewing, Noctuid moth, Reduviidae, Bombyliidae, Elateridae, Adalia decempunctata str Y, Halyzia*	Cream spot lady beetle, Green lace wings, owlet moths, kissing bugs, bee flies, click beetles, ten-spotted lady beetle, orange lady beetle, no common names	UK, Germany, Spain, New Zealand	none	none	no	(Weinert *et al.* 2009)

sedecimguttata str. *D, Scymnus suturalis* str. *theta, Subcoccinella vigintiquatuorpunctata strain J*	reported, 24-spot lady beetle PCR confirmed using 16s rRNA gene, *coxA*, *atpA*, and *gatB*					
Pnigalio soemius	Parasitoid wasp, Hymenoptera: Eulophidae PCR detection of *R. bellii*-like using 16s rRNA gene	Italy	*P. soemius*	*Methylobacterium* sp. and *Serratia proteamaculans*	not definitive	(Giorgini *et al.* 2010)
Macrosiphum euphorbiae	Potato aphid, Hemiptera: Aphididae PCR detection of *R. bellii*-like using 16s rRNA gene	France and Germany lab colonies	*M. euphorbiae*	*Buchnera aphidicola*	yes	(Francis *et al.* 2010)
Curculio sikkimensis	Chestnut weevil, Coleoptera: Curculionidae PCR detection of *R. bellii*-like using 16s rRNA gene	Japan	*C. sikkimensis*	*Sodalis*, *Wolbachia*, *Spiroplasm*	yes	(Toju and Fukatsu 2011)
Macrolophus pygmaeus	Hemiptera: Miridae PCR detection of *R. bellii*-like using 16s rRNA gene, *gltA*, and *coxA*	Greece, Spain, Italy	*M. pygmaeus*	*Wolbachia*, *R. limoniae*-like	potentially yes	(Machtelinckx *et al.* 2012)
Ceroplastes sinensis	Chinese white wax scale, Hemiptera: Coccoidea PCR detection of *R. bellii*-like using 16s rRNA gene	Yunan, Kunming, China	none	none	No	Xu, D.L. (at NCBI)

Nysius expresus, Nysius sp. 1 and 2	Hemiptera: Lygaeidae PCR detection of *R. bellii*-like 16s rRNA gene	Japan	*Nysius* sp. 1	*Wolbachia*	yes	(Matsuura *et al.* 2012)
Pnigalio soemius	Parasitoid wasp, Hymenoptera: Eulophidae PCR detection of *R. bellii*-like using 16s rRNA gene	Italy and Bulgaria	*P. soemius*	*Wolbachia*	potentially yes	(Gebiola *et al.* 2012)
Curculio kojimai/Koreoculio minuitissimus	Coleoptera: Curculionidae PCR detection of *R. bellii*-like using 16s rRNA gene	Japan, Indonesia	*C. kojimai/ K. minuitissimus*	*Curculionophilus* symbiont/ *Wolbachia*	yes/not definitive	(Toju *et al.* 2013)
Neophilaenus lineatus	Lined spittle bug, Hemiptera: Apophoridae PCR detection of *R. bellii*-like using 16s rRNA gene	Vermont	*N. lineatus*	*Sodalis, Sulcia*	not definitive	(Koga *et al.* 2013)
Hybos femoratus str. *247, Medetera abstrusa, Dolichopus longicornis* str. *158, Dolichopus linearis* str. *157, Chrysotus palustris str. 155, Dolichopus subpennatus* str. *153, Argyra perplexa* str. *120, Lamprochromus strobli* str. *152*	Diptera: Dolichopodidae and other Empoidea PCR detection of *R. bellii*-like using *gltA*	Belgium	*Hybos femoratus/ Medetera abstrusa/ Dolichopus longicornis/ Dolichopus linearis/ Chrysotus palustris/ Dolichopus subpennatus/ Argyra perplexa/ Lamprochromus strobli*	*Wolbachia, Serratia/ Serratia/ Serratia / Serratia / Serratia / Serratia /Wolbachia, Serratia / Serratia*	yes for all	(Martin *et al.* 2013)

Nesidiocoris tenuis (Reuter)	Tomato mirid, Hemiptera: Miridae PCR detection of *R. bellii*-like using 16s rRNA gene and *gltA*	Tomato field in Northeast Israel, and mass reared since 2010	*N. tenuis*	*Wolbachia*	potentially yes	(Caspi-Fluger *et al.* 2014)
Lissorhoptrous oryzophilus	Rice water weevil, Coleoptera: Curculionidae PCR detection of *R. bellii*-like using 16s rRNA gene	California, China	*L. oryzophilus*	*Wolbachia*	potentially yes	(Lu *et al.* 2014)
Aphis craccivora	Cowpea aphid, Hemiptera: Aphididae PCR detection of *R. bellii*-like using 16s rRNA gene	China, Serbia, Benin	*A. craccivora*	*Arsonophonus, Hamiltonella, Regiella, Serratia, Spiroplasma*	yes with Arsenophonus	(Brady *et al.* 2014)
Halictus tumulorum isolate *Hatu 16s*, *Lasioglossum zonulum* isolate *LaZo*, *Lasioglossum pauxillum* isolate *LP*, *Lasioglossum semilucens* isolate *LaSe*, *Halictus subauratus* isolate *HS*, *Lasioglossum laticeps* isolate *LaLa*	Hymenoptera: Anthophila PCR confirmed using 16s rRNA gene, *coxA*, *atpA*, and *gatB*	Germany	*Halictid*	*Wolbachia*	yes	(Gerth *et al.* 2015)

Genes used to confirm *R. bellii*-like agents were provided for future reference. The details of the geographical location and hosts were provided as accurately as described within the referenced articles.

Chapter 2: Structural modeling and transcriptional detection of the bacterial conjugation system of *Rickettsia bellii* RML 369-C

Introduction

Rickettsiae are obligate intracellular, Gram-negative bacteria found in nature coexisting with their arthropod hosts. The genomes of *Rickettsia* provide evidence of frequent gene loss, acquisition and, overall high turnover rate. *Rickettsia prowazekii* strain Madrid E, the first rickettsial genome to be sequenced, has a reduced genome correlating with the reductive genome of mitochondria (Andersson *et al.* 1998). Other rickettsial genomes also show evidence of a reductive evolutionary trend with gene loss leading to pathogenicity (Emelyanov 2001; Ogata *et al.* 2001; Malek *et al.* 2004; Mcleod *et al.* 2004; Blanc *et al.* 2007a; Gillespie *et al.* 2007, 2008; Felsheim, Kurtti and Munderloh 2009). Genes found in rickettsial genomes are frequently annotated as nonfunctional (due to premature stop codons or split genes), hypothetical, or unknown, taken as evidence for the reduction of rickettsial genomes and the acquisition of novel, unknown genes. Evidently, rickettsiae can acquire new genetic material (Ogata *et al.* 2006; Blanc *et al.* 2007a; Felsheim, Kurtti and Munderloh 2009; Gillespie *et al.* 2012, 2014) and turnover genes at a fast rate (Fuxelius *et al.* 2008; Merhej *et al.* 2011), so one can hypothesize that genes not undergoing reductive evolution are indispensible and beneficial. Therefore, these genes may be functional and biologically relevant, whether the acquired genes retain their original function or have been adapted for another function that improves the fitness of that rickettsia remains to be determined.

The genome of *Rickettsia bellii* RML 369-C contains acquired genes that are apparently functional (Ogata *et al.* 2006). Among these are the transfer (*tra*) genes that

are clustered together indicating an operon. This putative operon contains genes that code for proteins that constitute the machinery for bacterial conjugation consisting of two main components: mating pair formation and DNA transfer and replication. Mating pair formation consists of the pilus (*traA* in *E. coli*), the core complex (*traB, -V, -K, -N,* and *-F* in *E. coli*), the inner membrane platform (*traL, -G, -B* in *E. coli*), and the main ATPase responsible for the bioenergetics of pilus formation (*traC*) (Lawley *et al.* 2003; Cabezon *et al.* 2015). The DNA transfer and replication component comprises the relaxosome consisting of the relaxase (*traI* in *E. coli* and $traA_{Ti}$ in *R. bellii*), ATPase (*traD*), and accessory proteins (*traY, traM* and integration host factor in *E. coli*) (Firth, Ippen-Ihler and Skurray 1996; Cabezon *et al.* 2015). The function of bacterial conjugation is the transfer of genetic material from a donor to a recipient bacterium that usually confers some fitness advantage. Rickettsiae that have remnants or the complete *tra* cluster include but are not limited to *R. felis, R. massiliae, R. peacockii, R. parkeri,* and *R. buchneri* (rickettsial endosymbiont of *Ixodes scapularis*) (Ogata *et al.* 2005a, 2006; Blanc *et al.* 2007a; Felsheim, Kurtti and Munderloh 2009; Gillespie *et al.* 2012). Bioinformatics support a conclusion that these genes are related to the F-like and Ti-like system of bacterial conjugation. The F-like genes comprise genes functioning in the mating pair formation of bacterial conjugation, while the Ti-like system encodes the DNA replication and transfer proteins. Taken together, the *tra* genes of *R. bellii* 369-C apparently code for the essential components of bacterial conjugation.

To demonstrate that the *tra* genes of *R. bellii* are functional and have not undergone reduction and mutation, the mRNA transcripts and/or the proteins will have to be detected. In this chapter, I present evidence showing transcriptional activity of the *tra*

genes in *R. bellii*. The *tra* genes were shown to be transcriptionally active using reverse transcription-PCR (RT-PCR). The operon controls were also assessed indirectly using RT-PCR to detect the intergenic sequences of each *tra* genes. The prediction of the operons of the *tra* clusters of *R. bellii* RML 369-C and *R. massiliae* MTU5 were obtained from the Database for prOkaryotic OpeRon (DOOR) and compared with the operon of the bacterial conjugation genes of *Eshcerichia coli*. Furthermore, $traA_{Ti}$ transcripts were quantified relative to two housekeeping genes, *infB* (translation initiation factor IF-2) and *rpoB* (RNA polymerase subunit beta). Predictions of the transmembrane spanning domains and signal peptides of the *tra* genes indicate that rickettsial *tra* genes may form a conjugation structure distinctively from the F-T4SS model resulting in a different mechanism of bacterial conjugation. Yet, the *tra* genes may still retain the authentic function of transferring genetic material.

Materials and Methods

Generating rickettsial RNA

Rickettsia bellii was cultured in an *Ixodes scapularis* embryonic cell line, ISE6 (Munderloh *et al.* 1999), at 34^{o}C, 32^{o}C, and 26^{o}C and in African green monkey kidney cells, Vero (ATCC CCL-81), at 34^{o}C using a modified Leibovitz's L15 medium (L15C-300) supplemented with 10% fetal bovine serum (FBS), 5% tryptose phosphate broth, and 0.1% lipoprotein concentrate as described (Munderloh and Kurtti 1989; Oliver *et al.* 2013) in 25 cm^2 T-flasks (Greiner Bio-One, NC) for 72-96 hours until cytopathic effects, cell lysis and extracellular rickettsiae were observed in the host cell layer. Host cells and rickettsiae were released into suspension within the flask, transferred to 1.5 mL microcentrifuge tubes, and centrifuged at 13,600 x rcf for 5 minutes to collect host cells

and rickettsiae in pellets. The supernatant was discarded, cell/rickettsiae pellets originating from the same culture were combined, and TRI Reagent (Sigma, MO) was added to lyse the pellets. RNA was extracted according to the manufacturer's protocol with the exception of using heavy phase lock gel tubes (5 PRIME, MD) instead of standard microcentrifuge tubes. RNA samples were incubated with DNase turbo (ThermoFisher Scientific, NY) at 37^{o}C for 1 hour to digest contaminating DNA. RNA samples were stored at -80^{o}C until use.

Detecting transcripts of the *tra* genes

The transcripts of *traA*$_{Ti}$, *traH*, and *traF* were detected using the Access RT-PCR System (Promega, CA) and the primers listed in Table 2.1 using total RNA from infected cell cultures as templates. Additional RT-PCR was conducted for the intergenic sequences of the *tra* genes to identify potential operons. The primers and deoxynucleotide (dNTP) solution were used at a final concentration of 1 µM and the RNA template was at 100 µg per reaction. Thermal cycles were as follow: 1 cycle at 45^{o}C for 45 minutes for cDNA synthesis; initial denaturation at 94^{o}C for 2 minutes; 40 cycles DNA denaturation at 94^{o}C for 30 seconds, annealing at 46-50^{o}C for 1 minute, and extension at 68^{o}C for 1 minute; and, 1 cycle at 68^{o}C for 7 minutes for the final extension. The reactions were loaded onto a 1% agarose gel, electrophoresed at 80 V for 1.5 hours, and stained in 1x Gel Green (Biotium, CA) for 1 hour. The sizes of the RT-PCR products were determined using a 100 bp ladder (New England BioLabs, MA) although the ladder is not shown here. RT-PCR products were visualized by using a dark reader (Clare Chemical Research, CO) and an image was taken using ImageJ v. 1.23.

Predicting the *tra* operons of *R. bellii*

The gene synteny of the *tra* genes of *R. bellii* was generated using the genome data of *R. bellii* RML-369C (NC_007940.1) to show the relationship between each of the *tra* genes and the directons of transcription (immediate adjacent genes that are in the same direction). In addition, the prediction of the *tra* operons were obtained from DOOR, an operon prediction database (Mao *et al.* 2009, 2014). The data was extracted from DOOR and displayed in a table format. The operon of the *tra* genes of *E. coli* was compared to the predicted operons of *R. bellii* RML 369-C and *R. massiliae* MTU5 to identify similarities and differences in gene length, order (or synteny), and potential regulation.

Prediction of the transmembrane-spanning domains and signal peptide of the Tra proteins

SignalP v.3.1 (Bendtsen *et al.* 2004), TMHMM v.2.0 (Sonnhammer, von Heijne and Krogh 1998; Krogh *et al.* 2001), Phobius (Kall, Krogh and Sonnhammer 2004, 2007), and LipoP v.1.0 (Juncker *et al.* 2003) were used to predict transmembrane domains and signal peptides of the *tra* genes. Additionally, BOCTOPUS (Hayat and Elofsson 2012) was used to predict beta barrel structures of *traN*. Changes to the currently modeled putative conjugation system of *Ricketsia* were based on the collective information of the predicted domains and current information on the F-T4SS structure (Firth, Ippen-Ihler and Skurray 1996; Lawley *et al.* 2003; Cabezon *et al.* 2015; Gillespie *et al.* 2015).

Relative transcription of $traA_{Ti}$ to two housekeeping genes: *infB* and *rpoB*

A one-step quantitative reverse transcriptase PCR (qRT-PCR) using total RNA from rickettsia-infected host cells at [3.2 ng/μL], the Brilliant II SYBR Green QRT-PCR

Master Mix kit (Agilent, CA), and primers listed in Table 2.1 at [120 nM] was used to quantify the number of mRNA copies of *traA*$_{Ti}$, *infB* and *rpoB*. Thermal cycles are as follow: 1 cycle at 50^{o}C for 30 minutes for cDNA synthesis; 1 cycle at 95^{o}C for 10 minutes; 40 cycles at 95^{o}C for 30 seconds, 48^{o}C for 1 minute, 68^{o}C for 1 minute; and a dissociation curve of 1 cycle at 95^{o}C for 1 minute, 50^{o}C for 30 seconds, and 95^{o}C for 30 seconds to evaluate a single product for each reactions. Fold changes of *traA*$_{Ti}$ to *infB* and *rpoB* were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001). SigmaPlot v. 13 (Systat Software, Inc., CA) was used to generate a line graph and a one-way ANOVA was run for statistical analysis using the Holm-Sidak correction with an alpha set at 0.05.

Results

Transcriptional detection of the *traA*$_{Ti}$, *-H*, and *-F*

RT-PCR was used to detect transcription of the *tra* genes of *R. bellii*. *traA*$_{Ti}$, *traH*, and *traF* were transcribed in *R. bellii* grown in ISE6 cells cultured at 26^{o}C, 32^{o}C, and 34^{o}C (Fig. 2.1A). The detection of *traA*$_{Ti}$ transcripts in *R. bellii* cultured in Vero cells at 34^{o}C suggests the detection of the rest of the *tra* genes (Fig. 2.1B). The different cell types and temperature at which the *tra* genes were detected implies that the *tra* genes can remain active in variable conditions that favor the transfer of genetic material. This implies that the temperature at which host cells are maintained and the different host cells do not affect the transcription of the *tra* genes of *R. bellii*.

The *tra* genes are grouped into multiple operons

Operons are usually predicted based on the direction of adjacent genes, intergenic distances, conserved orthologous gene pairs and neighborhood, relatedness of adjacent gene function, correlating transcription numbers of adjacent genes, ratio of the gene

length and length of adjacent genes, and frequency of DNA motifs within intergenic sequences (Dam *et al.* 2007; Taboada, Verde and Merino 2010; Taboada *et al.* 2012). Here, the intergenic sequences of the *tra* genes in *R. bellii* were detected using RT-PCR from the first end of the *tra* gene cluster that consists of the intergenic sequence of RBE_0439, a hypothetical gene predicted to encode a protein rich in leucine repeats, and *traL* to the intergenic sequence of $traD_{Ti}$ and RBE_0419, a predicted guanosine polyphosphate pyrophosphohydrolase, which is the other end of the *tra* gene cluster in *R. bellii* (Fig. 2.2). A large directon (adjacent genes that share the same orientation and direction) is located from RBE_0442 to $traD_F$ with a bicistronic directon of $traA_{Ti}$ and RBE_0422 (Fig. 2.2, Table 2.3), and a monocistronic directon of $traD_{Ti}$ (Fig. 2.2, Table 2.3). Interestingly, the intergenic sequences of neighboring genes oriented differently (i.e. $traD_F$ and RBE_0422, $traA_{Ti}$ and $traD_{Ti}$) were still detected (Fig. 2.3) along with the rest of the intergenic sequences detected using primers in Table 2.1. In a permissive and nutrient rich environment, termination of transcripts may not be stringently regulated causing the detection of convergent genes when using RT-PCR (Woodard and Wood 2011).

In silico prediction of non-canonical prokaryotes may provide clues to identify how the *tra* genes are grouped into operons. The predicted operons of the *tra* genes of *R. bellii* and *R. massiliae* were found in DOOR (Dam *et al.* 2007; Mao *et al.* 2009, 2014) and ProOpDB (Taboada, Verde and Merino 2010; Taboada *et al.* 2012). The results from ProOpDB predicted monocistronic operons that clearly contradict the overlapping adjacent genes of *trbC* to *traH* of *R. massiliae* that is a hallmark of an operon and will not be discussed further. The operon predictions from DOOR for the *tra* genes of *R. bellii*

and *R. massiliae* were compared to the operon in *E. coli* (Table 2.3). The F-like *tra* genes of *R. bellii* and *R. massiliae* are oriented in the same direction and sequentially ordered similar to each other and to the *tra* cluster of *E. coli*. Regulation of the *tra* genes is fundamentally different, based on predictions of multiple operons and directons in *R. bellii* and *R. massiliae* when compared to a single operon for all the *tra* genes in *E. coli*. Furthermore, the orientation of the F-like *tra* genes of both *R. bellii* and *R. massiliae* is similar to the directon of the F plasmid of *E. coli* indicating gene integrity and possibly functionality because of the conserved genes and gene orientations.

The predicted relaxase, $TraA_{Ti}$, may illuminate when bacterial conjugation occurs

If the *tra* genes are functional, then the timing of the transcription of the *tra* genes may determine when the phenotype of the *tra* genes can be observed. $traA_{Ti}$ encodes the relaxase of bacterial conjugation and is responsible for the DNA transfer and replication processes. Nicking and generating the transfer DNA by $TraA_{Ti}$ during conjugation should provide evidence as to when bacterial conjugation is happening. The relative transcripts of $traA_{Ti}$ to *infB* and *rpoB* were determined using qRT-PCR. *traATi*:*rpoB* at 96 HPI showed a significant increase compared to 24, 48, and 72 HPI while *traATi*:*infB* display a significant increase at 96 HPI when compared to 48 and 72 HPI, but not to 24 HPI (Fig. 2.4). Although neither housekeeping gene showed significant differences at each time point, the general trend of $traA_{Ti}$ to both *infB* and *rpoB* still indicated increased relative transcription at 96 HPI (Fig. 2.4).

Revising the localization of TraW and TraH in the bacterial conjugation system of rickettsiae

Structure of the mating-pair formation

The Tra proteins were modeled according to the F-like bacterial conjugation type IV secretion system (F-T4SS) (Firth, Ippen-Ihler and Skurray 1996; Lawley *et al.* 2003; Cabezon *et al.* 2015) and based on the consensus of the predicted transmembrane domains and signal peptide sequence. The general components of the F-T4SS consists of the mating pair formation proteins, which comprises the F-pilus, the core complex, the inner membrane platform, and the cytoplasmic ATPases (Cabezon *et al.* 2015), and the DNA transfer and replication components specific for genetic transfer. The primary constituent of the pilus is the $TraA_F$ pilin protein. The $traA_F$ pilin gene is yet to be identified in *R. bellii* but has possibly been annotated as the C-terminal of *R. bellii traK,* which contains a signal peptide and two transmembrane domains that span the inner membrane (all predicted signal peptide and transmembrane domains can be found in Table 2.2) (Lawley *et al.* 2003; Gillespie *et al.* 2009; Heu *et al.* 2015).

F-T4SS specific proteins of R. bellii

traL, -E, -K, -B, -V, -W, -U, -F, -H, -G, trbC, and *trbI* affects the morphology and assembly of the pilus and mutation of any of these genes prevents bacterial conjugation (Firth, Ippen-Ihler and Skurray 1996). All, except for *trbI*, have been annotated in *R. bellii* (Ogata *et al.* 2006). *traL, -E, -K, -B, -V,* and -G of *R. bellii* will be discussed later along with the structural and functional components. TraU and TrbC of *R. bellii* are both predicted to have a signal peptide at the N-terminus by Phobius and SignalP and Phobius, respectively, that localizes the proteins to the periplasm as found similarly in F-T4SS (Lawley *et al.* 2003). TraF and TraH of *R. bellii* had contradictory results with a transmembrane domain predicted in TMHMM and a signal peptide predicted in Phobius. In F-T4SS, TraF and TraH are localized in the periplasm, but form an outer membrane

protein complex directly or indirectly with TraV, respectively (Arutyunov *et al.* 2010). Furthermore, TraH apparently interacts directly with TraF and TraU, but indirectly with TraW through TraU in the F-T4SS (Harris and Silverman 2004; Arutyunov *et al.* 2010). TraF and TraH of *R. bellii* will be modeled as periplasmic proteins with a caveat that both proteins may be integrated into the inner or outer membrane. TraW of *R. bellii* was predicted to have a transmembrane-spanning domain by TMHMM and Phobius; thus, contradicting the periplasmic nature of TraW of the F-T4SS (Maneewannakul, Maneewannakul and Ippen-Ihler 1992). In the *Rickettsia* model, TraW will be integrated into the inner membrane and hypothesized to anchor the protein complex of TraV, -F, -U, -W, and possibly TrbC and TraH onto the inner membrane. Therefore, TraF and –H will be shown as periplasmic proteins with a caveat of integrating into the inner or outer membrane and TraW is predicted to integrate into the inner membrane.

The core complex of F-T4SS of R. bellii

The core complex consists of the envelope spanning protein complex of *traB*, *-K*, and *–V* in *E. coli* (Harris, Hombs and Silverman 2001; Lawley *et al.* 2003) and similarly modeled with the same genes in *R. bellii* except *traK* is split into two separate genes: $traK_C$ and $traK_N$ for the carboxyl-terminus and amino-terminus, respectively. $TraK_C$ and $TraK_N$ may interact with each other given both genes are part of a complete gene found in *E. coli*, with $TraK_C$ containing a signal peptide and two transmembrane domains, as predicted by TMHMM and Phobius, that spans the outer membrane. The interaction of $TraK_C$ and $TraK_N$ has to precede the cleavage of the signal peptide because $TraK_N$ lacks a predicted signal peptide allowing, hypothetically, $TraK_N$ to exploit the signal peptide of $TraK_C$ when $TraK_C$ gets directed to the periplasm. In the F-T4SS, the interaction of TraK

and TraV has been shown to occur in the outer membrane where TraV is a lipoprotein that anchors to the outer membrane and this may be true for $TraK_N$, $TraK_C$ and TraV of *R. bellii* (Doran *et al.* 1994; Lawley *et al.* 2003). LipoP analysis of *traV* of *R. bellii* shows a predicted signal peptide that is cleaved by signal peptidase II immediately before the cysteine at position 21, both are characteristics of lipoproteins (Nakayama, Kurokawa and Lee 2012). The + 2 amino acid of the predicted mature TraV protein is a lysine and not an aspartate and this supports the localization of TraV to the inner leaflet of the outer membrane of *R. bellii*. Similarly to TraV of *E. coli*, the following two cysteines at position 12 and 20 of TraV of *R. bellii* may affect the transport of TraV into the outer membrane (Harris and Silverman 2002). Collectively, this evidence suggests that TraV of *R. bellii* is a lipoprotein that localizes to the outer membrane and provides an outermembrane anchor to the core complex of TraB, -K, and -V. In addition, TraB of *R. bellii* is predicted to have a transmembrane in the N-terminus by TMHMM and contradicted by Phobius to contain a signal peptide. However, the TraB of *E. coli* was proven to fractionate with the inner membrane and to contain a predicted transmembrane domain at the N-terminus by TMHMM and Phobius. Furthermore, TraB and its homologues shares a common transmembrane domain at the N-terminus (Lawley *et al.* 2003). Therefore, TraB of *R. bellii* was also predicted to have a transmembrane domain that anchors the TraB, -K, and –V protein complex to the inner membrane. Overall, the core complex is similar to the F-T4SS.

The inner membrane platform of the F-T4SS of R. bellii

The inner membrane platform is hypothesized to consist of TraG, TraE, and TraL where TraE and TraL are proteins that have been proposed to stabilize and interact with

other components of the F-T4SS (Guglielmini *et al.* 2014). TraE and TraL of *R. bellii* were predicted to contain transmembrane-spanning domains by TMHMM and Phobius. This corroborates with the evidence of TraE and TraL integrating into the inner membrane of *E. coli* in the F-T4SS that functions to help with the pilus tip formation (Firth, Ippen-Ihler and Skurray 1996; Lawley *et al.* 2003). TraG with TraE and –L forms the inner membrane platform in the F-T4SS. The N-terminus of TraG is homologous to VirB6 in the *virB/D* system, which localizes to the inner membrane and within the periphery of the core complex (Cabezon *et al.* 2015). The predicted multiple transmembrane domains in the N-terminus (3 by TMHMM, 4 by Phobius) that span approximately the first half of TraG of *R. bellii* supports the integration into the inner membrane. The C-terminus of TraG is periplasmic and functions in mating-pair stabilization and has been proposed to interact with the other mating-pair stabilization protein, TraN, in the F-T4SS system (Lawley *et al.* 2003). The whole TraG protein functions in entry exclusion of the transfer DNA as a mechanism to avoid donor-to-donor conjugation. Therefore, the proposed function of TraG and the predicted multiple transmembrane domains in the N-terminus collectively support the hypothesized localization of TraG in the inner membrane of *R. bellii*.

The cytoplasmic ATPases of the F-T4SS of R. bellii

The cytoplasmic ATPases are encoded by *traC* and $traD_F$ of *R. bellii* and are represented by the well-studied *virB4* and *virD4* genes, respectively. TraC and TraD are hexameric and empirical evidence of nucleotide hydrolysis in homologous proteins has been observed, e.g. TrwK and TrwB of the plasmid R388, respectively (Tato *et al.* 2007; Arechaga *et al.* 2008; Cabezon *et al.* 2015). TraC is a cytoplasmic component that

attaches to the inner membrane and/or core complex via TraL and TraB, respectively, and generates energy for pilus assembly. TraC is also hypothesized to warp the core complex during pilus assembly and ATP hydrolysis in the F-T4SS (Lawley *et al.* 2003; Cabezon *et al.* 2015). The absence of a signal peptide or transmembrane domain predicted by SignalP, TMHMM, and Phobius supports TraC of *R. bellii* as a cytoplasmic protein. The other cytoplasmic ATPase is $TraD_F$ of *R. bellii*. Predicted transmembrane domains (2 in TMHMM and 3 in Phobius) at the N-terminus are hypothesized to directly anchor $TraD_F$ to the inner membrane and help stabilize the protein, similarly in the F-T4SS (Cabezon *et al.* 2015). $TraD_F$ may function in coupling the relaxosome to the transferosome (or mating pair formation structure) by interacting with TraB as demonstrated with the interaction of the homologues TrwB and TrwE, respectively (Llosa, Zunzunegui and de la Cruz 2003). The cytoplasmic ATPases are usually conserved within T4SS, so the predictions of these proteins and the suspected subcellular localization of TraC and $TraD_F$ of *R. bellii* were not expected to deviate from the general T4SS model.

DNA transfer and replication

The DNA transfer and replication component of the F-T4SS consists of the relaxosome and the coupling protein (TraD). Notice that the TraD is part of both components of the F-T4SS and have already been discussed previously (Fig. 2.1B). In the model for *R. bellii*, the predicted main enzyme responsible for the DNA transfer and replication is the relaxase $TraA_{Ti}$. This enzyme is related to the relaxase of *Agrobacterium tumefaciens*; thus, making the bacterial conjugation of *R. bellii* a chimera of both the *tra* genes of the F plasmid and *A. tumefaciens* (Ogata *et al.* 2006). The function of the relaxase of the F-T4SS, TraI, can still be applied to $TraA_{Ti}$ as both share

domains that function similarly by recognizing a sequence called the oriT (origin of transfer), nicking the DNA on one strand, and helically unwinding the plasmid to generate a transfer strand in a rolling circle replication fashion (Alvarez-Martinez and Christie 2009). Interaction of the relaxase with the coupling protein, $TraD_F$, initiates the nicking of oriT followed by the transfer of the relaxase covalently bound to the transfer-DNA through the transferosome to the recipient. The prediction programs did not detect a signal peptide or any transmembrane-spanning domain in $TraA_{Ti}$. Thus, this lack of detection supports $TraA_{Ti}$ to be a cytoplasmic protein that corroborates with the predicted function of catalyzing the transfer DNA.

Discussion

The genome of *Rickettsia bellii* RML 369-C has the complete set of *tra* genes predicted to function in bacterial conjugation. The integrity and activity of the *tra* genes have not been tested. In this experiment, the *tra* genes were shown to be transcriptionally active (Fig. 2.1). Transcripts spanning the intergenic sequences of the *tra* genes were also detected despite the converging and diverging adjacent genes (Fig. 2.2). *Rickettsia prowazekii* also demonstrated the lack of transcription termination of neighboring genes that converged together with a few convergent gene pairs showing a longer transcript on one strand compared to the other sense strand (Woodard and Wood 2011). Thus, detecting the intergenic sequence without a sense strand orientation in *Rickettsia* may not identify potential transcriptional promoters. Further analysis will be needed to include probes specific for differentiating the sense and antisense strand of the intergenic sequence to distinguish the regulation of transcription termination. However, the *tra* gene operon could still be evaluated based on prediction programs. DOOR has predicted the

tra genes of *R. bellii* to exist as multiple operons and was able to predict the *tra* genes that show nucleotide overlap as operons. In contrast, ProOpDB predicted variable operon lengths for the overlapping adjacent genes of *trbC* to *traH* of *R. massiliae*, but was able to predict a monocistronic operon of the same genes in *R. bellii*. Overall, the *tra* genes of *R. bellii* are proposed to exist as multiple operons.

Rickettsiae may have grouped the F-like *tra* region into multiple operons by reducing and deleting *tra* genes of the F-T4SS that may have not been essential in bacterial conjugation. For example, TraX of F-T4SS has been shown to be dispensable in bacterial conjugation, yet it is maintained within the F-T4SS in *E. coli* (Maneewannakul, Maneewannakul and Ippen-Ihler 1992). Evidence of reductive genome evolution in rickettsiae is prevalent based on genomic comparisons to their free-living counterparts and likely argues for the reduction in dispensable genes. Therefore, the *tra* genes of *R. bellii* still maintain some synteny with the *tra* genes of the F-T4SS of *E. coli*, but the *tra* genes of *R. bellii* appear to have kept the minimal *tra* genes necessary to perform bacterial conjugation.

The *tra* genes are maintained and actively transcribed within *R. bellii*, but when will the *tra* genes be the most active? This study uses two housekeeping genes as reference genes to quantify the relative number of transcripts of an essential enzyme of bacterial conjugation- $TraA_{Ti}$. Up-regulation of this gene was shown to occur at 72 HPI compared with 12 HPI when using *rpoB* as a reference gene. However, when using *infB* as a reference gene, $traA_{Ti}$ did not show any significant difference at 72 HPI compared to 12 HPI. Further detailed analysis should include the validation of these reference genes to accurately depict relative transcription of $traA_{Ti}$. In addition, the rest of the *tra* genes

should be quantified to show activity at the later exponential or early stationary growth phase.

Although the quantification of the transcripts may elucidate the timing and growth phase of bacterial conjugation, the predicted structure should also be modeled to observe for potentially similar structures as the F-T4SS of *E. coli*. The predicted transmembrane-spanning domains and signal peptides of the constituents of the bacterial conjugation machinery of *R. bellii* were shown to align with the *E. coli* F-T4SS overall (Table 2.2). TraF and TraH of *R. bellii*; however, were predicted to have contradicting domains and one should be cautious in placing these two proteins as periplasmic or integral proteins. The main difference is the localization of TraW of *R. bellii* because both TMHMM and Phobius predicted a transmembrane domain instead of a signal peptide as seen in the prediction of TraW of *E. coli*. It could be that TraW is integrated within the inner membrane of *R. bellii* and anchors part of the outer membrane protein complex. It is important to note that current knowledge of this system has yet to identify other components of the *tra* system in *R. bellii*. This model only depicts what is predicted to be the complete bacterial conjugation complex.

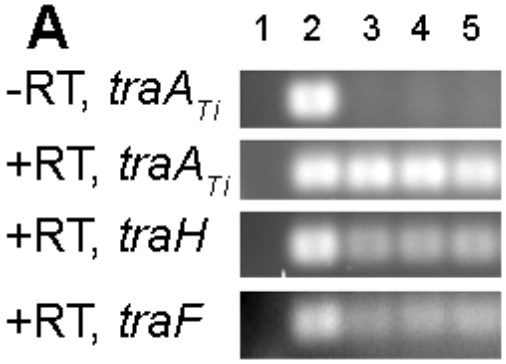
Among the *tra* genes of *R. bellii* and *R. massiliae* there are several other genes, known as *tra*-associated genes that were omitted from Table 2.3 to simplify the analysis of the *tra* genes, that may have piggybacked with the *tra* genes (Ogata *et al.* 2006; Blanc *et al.* 2007a). Differences in these genes suggest different origins of the *tra* genes even though the *tra* genes are similar via gene order, orientation, and nucleotide sequences. Furthermore, different copies of the *tra* cluster in the genome of *R. buchneri* suggest multiple invasions and origins of the *tra* cluster. As shown by phylogenetic analysis, the

rickettsial amplified genetic element (RAGE)-RBE of *R. buchneri* was closest to the *tra* genes in *R. bellii* RML 369-C and the RAGE-B was closest to the *tra* genes found in *R. massiliae* MTU5 (Gillespie *et al.* 2012). A comparison between the complete genomes of different *R. massiliae* strains clearly shows that the *tra* cluster was recently acquired. *Rickettsia massiliae* strain MTU5 has the *tra* cluster while strain AZT80 did not. This suggests that *R. massiliae* MTU5 and perhaps other rickettsiae may have acquired the *tra* genes recently. Although *Rickettsia peacockii* also maintains a few remnants of the *tra* genes, the genomic shuffling of *R. peacockii* by homologous recombination at multiple transposon locations demonstrates that this process occurred after acquiring the *tra* genes (Felsheim, Kurtti and Munderloh 2009). It could be that the acquisition of the *tra* genes favors the incorporation and invasion of other mobile genetic material that eventually led to the genome shuffling of *R. peacockii*. Thus, the *tra* genes may contribute to the plasticity of rickettsial genome preceding reductive genome evolution.

Conclusion

In conclusion, the *tra* genes are transcriptionally active in ISE6 and Vero cell cultures and at different temperatures. The transcriptional regulation of the *tra* genes, and perhaps most genes, is loosely regulated similarly to *R. prowazekii* (Woodard and Wood 2011). Furthermore, *in silico* predictions group the *tra* genes into multiple operons differentiating it from the *tra* genes of *E. coli*. Future studies should include a validation of the reference gene for relative quantification and the quantification of all the *tra* genes to identify when all the *tra* genes are transcriptionally active and verify the predictions of the *tra* gene operons.

Fig. 2.1. Detection of the transcripts of the *tra* genes of *R. bellii*. RT-PCR was used to detect the transcripts of the *tra* genes using primer pairs listed in Table 2.1. A) $traA_{Ti}$, *traH*, and *traF* were detected in the total RNA samples of *R. bellii*-infected ISE6 cells. Lanes: 1 = H_2O (negative control), 2 = DNA of *R. bellii* (positive control), 3 = total RNA of *R. bellii*-infected ISE6 cultured at 34°C, 4 = 32°C, and 5 = 26°C. B) Transcripts of *gltA* were detected using CS-239/ CS-1069 primer pair to confirm the presence of *R. bellii* from total RNA of *R. bellii*-infected Vero cells. Subsequently, $traA_{Ti}$ was detected in total RNA samples of *R. bellii*-infected Vero cells. Lanes: 1 = H_2O (negative control), 2 = DNA of *R. bellii* (positive control), 3 = no RT with total RNA of *R. bellii*-infected Vero cells, and 4 = + RT with total RNA of *R. bellii*-infected Vero cells.



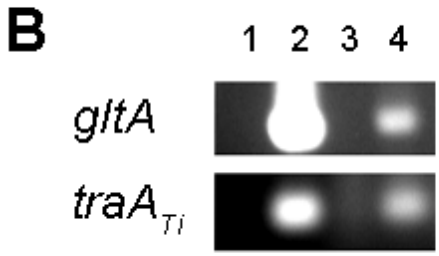


Fig. 2.2. Organization of the *tra* genes of *R. bellii*. An arrow represents a *tra* gene and its direction. Black arrow represents a gene that has at least one nucleotide in between adjacent genes. Arrows with patterns indicate genes that overlap or have no nucleotide between adjacent genes as indicated by the key. Genes are not to scale. Rec = recombinase, Int = integrase, Lrp = leucine-rich repeat protein, L = *traL*, E = *traE*, Kn = amino-terminus of *traK*, Kc = carboxyl-terminus of *traK*, B = *traB*, V = *traV*, C = *traC*, W = *traW*, Phd = prevent-host-death protein, U = *traU*, N = *traN*, F = *traF*, H = *traH*, G = *traG*, Tpr = tetratricopeptide repeat containing protein, D_F = $traD_F$ from *E. coli traD* system, Tnp = transposase, A_{Ti} = $traA_{Ti}$ from *A. tumefaciens* Ti plasmid, D_{Ti} = $traD_{Ti}$ from *A. tumefaciens* Ti plasmid, and SpoT16 = guanosine polyphosphate pyrophosphohydrolase.

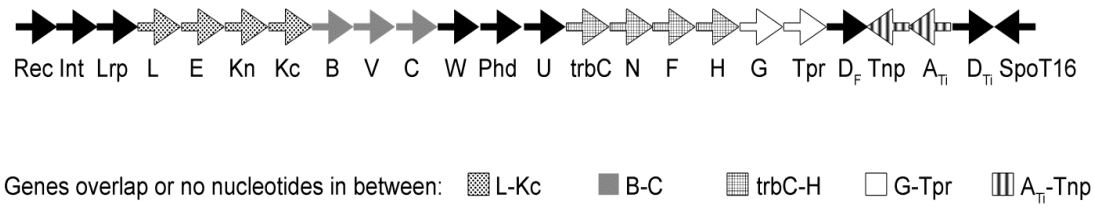


Fig. 2.3. Detecting intergenic sequences of the *tra* genes of *R. bellii*. RT-PCR was used to detect the intergenic sequences of the *tra* genes using primers listed in Table 2.1. Lanes are grouped into three lanes for each primer pair and starting from left to right is H_2O (negative control), DNA of *R. bellii* (positive control), and total RNA of *R. bellii*. Primer pairs used from left to right with lane numbers in parenthesis are as follow: RBE_0439-0438 or *traL* (1, 2, 3); RBE_0438 or *traL- traE* (4, 5, 6); *traE*-RBE_0436 (7, 8, 9); RBE_0436-0435 (10, 11, 12); RBE_0435-*traB* (13, 14, 15); *traB-traC* (16, 17, 18); *traC-traW* (19, 20, 21); *traW-traU* (22, 23, 24); *traU-trbC* (25, 26, 27); *trbC-traN* (28, 29, 30); *traN-traF* (31, 32, 33); *traF-traH* (34, 35, 36); *traH-traG* (37, 38, 39); *traG*-RBE_0424 (40, 41, 42); RBE_0424-$traD_F$ (43, 44, 45); $traD_F$-RBE_0422 (46, 47, 48); RBE_0422- $traA_{Ti}$ (49, 50, 51); $traA_{Ti}$-$traD_{Ti}$ (52, 53, 54); $traD_{Ti}$ -*spoT16* (55, 56, 57). Product size is indicated in Table 2.1. Arrows above lanes 46-48 and 52-54 indicate the direction of the adjacent genes for which the intergenic sequences were detected. Detection of the intergenic sequences of the converging and diverging adjacent genes indicates a lack of transcriptional termination.

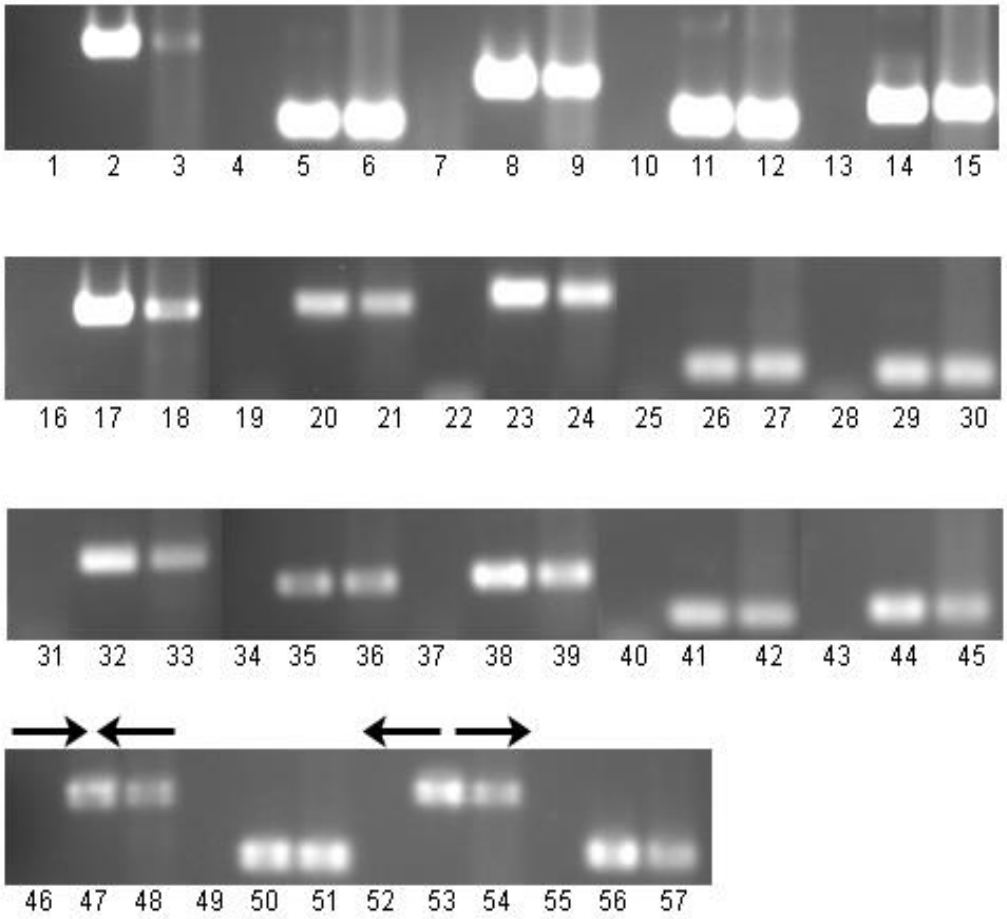

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
46 47 48 49 50 51 52 53 54 55 56 57

Fig. 2.4. An up-regulation trend of *traA$_{Ti}$*. The transcriptions of *traA$_{Ti}$*, *infB*, and *rpoB* were detected using qRT-PCR from total RNA of *R. bellii*-infected ISE6 at different time points. Statistical significance was evaluated with an alpha set at 0.05 and depicted using lower case and upper case letters for *traA$_{Ti}$*: *rpoB* and *traA$_{Ti}$*:*infB* fold changes, respectively.

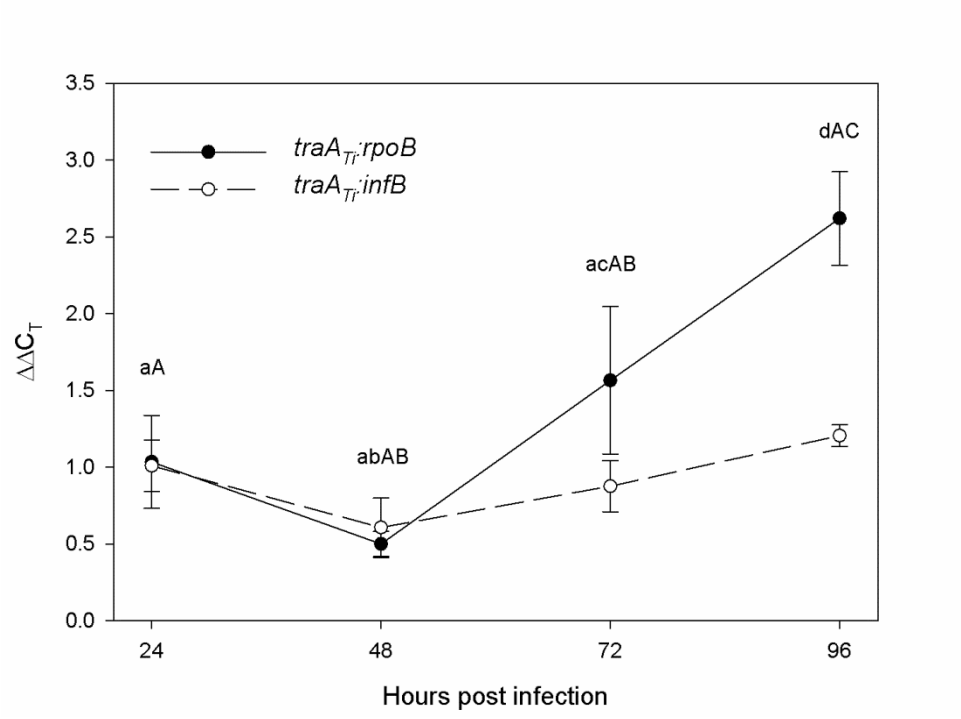


Table 2.1. *tra* PCR primers and sequences

Primers	Sequence	Product size	Source
TraA bellii F	TGGCACAGCAGAAAATATCG	218	Weinert et al. 2009
TraA bellii R	GAGATGGCTTTTGCCTTGAA		
TraH bellii F	TTAAGCCCTCTCTTTAGCGA	198	Weinert et al. 2009
TraH bellii R	CCTCAGAGGTAAGAAAAGCA		
TraF bellii F	ACAATAGGGATAAATGCCTGG	103	This study
TraF bellii R	GATAAAGACAGCAGCCAAAAAC		
CS-239	GCTCTTCTCATCCTATGGCTATTAT	830	Labruna et al. 2004
CS-1069	CAGGGTCTTCGTGCATTTCTT		
Rbe_0439-0438 F	ATCCCAAGCTTAGCATGTCC	675	This study
Rbe_0439-0438 R	AATGAACAGCTTGAACCGAA		
Rbe_0438- traE F	TCCGACGATCTGGCTCTATT	338	This study
Rbe_0438- traE R	GCTAAGCTTGGGATGATATGC		
traE-Rbe_0436 F	AACCTGAATTATGCACTGCT	507	This study
traE-Rbe_0436 R	AGCAGATATGAAGGTGGTGT		
Rbe_0436-0435 F	AGCCTTAAATTACCCCTGAC	376	This study
Rbe_0436-0435 R	ATGGTACGGTCCAAGATTTA		
Rbe_0435-traB F	CCTGATATTTCCCTTGATTG	457	This study
Rbe_0435-traB R	TAGAAGGACAGTTAGACCG		
traB-traC F	GCCAAGCAGGACAAATCCTACTGA	675	This study
traB-traC R	ACGAGCAGAAAGTATGAGTCCGGT		
traC-traW F	TGGAAATACATGACCTTGGACTTC	499	This study
traC-traW R	TGATCGCAGGAATGAGCGTTACAG		
traW-traU F	AGCCATAGGGCTTGGCACTACTTT	506	This study
traW-traU R	AGGTGGCGTGCTTACTACTCGTTT		

traU-trbC F	GTTGTGCCAAAGAACTTGTTCCA	219	This study
traU-trbC R	TGGTGAGACTGAAACTACCTGGCA		
trbC-traN F	CTGTGTTCATACTGTCTGGCACT	205	This study
trbC-traN R	GTAGTGATAGATAATAACCAAGGC		
traN-traF F	AACCTTTTGGTTGAGTGCTA	435	This study
traN-traF R	CTATTAGTTTTGGCACAGGG		
traF-traH F	CCTGCTGCATAATAATAGCC	351	This study
traF-traH R	AACACTAAATCCTGCAAAGC		
traH-traG F	CAAAGACTCTTCCTATGCCA	355	This study
traH-traG R	CCTCAGAGGTAAGAAAAGCA		
traG-Rbe_0424 F	GCTTTGTTGCTCGATTTGTTG	232	This study
traG-Rbe_0424 R	CAGAACTGGCCAGCAAATAGC		
Rbe_0424-traD F	GTACGCCACATTAGCCAAACTACAG	200	This study
Rbe_0424-traD R	GGTAATGATCTTATGGCTGACGAGT		
traD-Rbe_0422 F2	GTAACCGGTCTTGATGAAGC	717	This study
traD-Rbe_0422 R2	TATGGTGCTAACTCCATGCG		
Rbe_0422-traAti F	AAGCCATCTCGTCCCTGAAG	471	This study
Rbe_0422-traAti R	CGACAAGCAAAAATAAGTATCCGTC		
traAti-traDti F	CAAGTTTAGCTTTAACAACTAAGCC	748	This study
traAti-traDti R	GCAAGCATTTCCTCCTGAAC		
traDti-spoT16 F	GGACTATGTCACAACTCCTTGC	500	This study
traDti-spoT16 R	TTGAAATCGGCGGCTTAG		
Rbe369C rpoB F2	TGATGCCCAACCAAGATGAG	245	This study
Rbe369C rpoB R2	CAAAGCGGAGACGATTTACC		
Rbe369C infB F1	ACGCTACCTGAAGTTATTGGTG	189	This study
Rbe369C infB R1	CAACATCCGATTCTTGGACTC		

Table 2.2. Predicted signal peptides and transmembrane domains

R. bellii 369-C/ *E. coli* F plasmid	SignalP v.4.1	TMHMM v.2.0	Phobius	LipoP v.1.0
NA/ *traA*	(-/-)	(-/i31-53o, o77-94i, i101-120o)	(-/1-51sp, o75-94i, i101-120o)	(-/-)
traL/ traL	(-/-)	i20-42o, o47-64i/o38-60i)	(i20-40o, o46-64i/o38-60i)	(-/-)
traE/ traE	(-/-)	(i20-39o/i12-33o)	(i20-39o/i12-33o)	(-/-)
(traKn/traKc)/ traK	(-,-/-)	((o10-29i/i20-42o, o57-79i, i86-103o)/i7-29o)	((1-28sp/o20-42i, i63-79o, o85-103i)/1-21sp)	(-/21-22spI)
traB/ traB	(-/-)	(-/i13-32o)	(i40-58o/i12-33o)	(-/28-29spI)
traV/ traV	(19/-)	(-/-)	(1-21sp/1-21 sp)	(20-21spII/18-19spII)
traC/ traC	(-/-)	(-/-)	(-/o665-683i)	(-/-)
traW/ traW	(-/18)	(i7-29o/-)	(i7-26o/1-18sp)	(-/-)
traU/ traU	(-/22)	(-/-)	(1-18sp/1-22sp)	(-/22-23spI)
trbC/ trbC	(23/21)	(-/-)	(1-23sp/1-21sp)	(23-24 spI/21-22spI)
traN/ traN	(22/18)	(-/-)	(12-22sp/1-18sp)	(-/18-19spI)
traF/ traF	(-/29)	(i5-27o/-)	(1-36sp/1-29sp)	(-/29-30spI)
traH/ traH	(-/24)	(i12-34o/-)	(1-33sp/1-24sp)	(-/24-25spI)
traG/ traG	(-/-)	(i56-78o, o344-366i, i379-401o/o327-349i, i356-378o	(o58-78i, i348-373o, o379-402i, i423-446o/o32-50i, i57-73o, o331-349i, i361-385o)	(-/-)
traD/ traD	(-/-)	(i21-43o, o123-145i/o25-47i, i108-130o)/1-21sp)	(i21-43o, o124-145i, i389-407o/i28-47o, o108-134i)	(-/-)
traATi/ traI	(-/-)	(-/-)	(-/-)	(-/-)
traDTi/ NA	(-/-)	(-/-)	(-/-)	(-/-)

SignalP: numbers indicate the amino acid position where signal peptidase is predicted to cleave after. TMHMM: i = inside, o = outside or cytoplasmic/non-cytoplasmic, respectively. Sp = signal peptide predicted. LipoP: spI = signal peptidase I, spII = signal peptidase II.

Table 2.3. Comparative analysis of the *tra* operons of *R. bellii*, *R. massiliae*, and *E. coli*

***E. coli* str. 042 pAA**		***R. bellii* RML369-C**		***R. massiliae* MTU5**	
Gene	*Size (bp)*	*Gene*	*Size (bp)*	*Gene*	*Size (bp)*
traM	384	not determined		not determined	
$traJ^{B}$	690	not determined		not determined	
traY	396	not determined		not determined	
traA	366	not determined		not determined	
traL	312	$traL^{B}$	309	$traL^{B}$	315
traE	567	traE	552	traE	555
traK	729	not determined		not determined	
traB	1428	traB	1419	traB	1413
traV	516	traV	255	traV	240
traC	2628	$traC^{E}$	2589	$traC^{E}$	2520
traW	633	traW	765	traW pseudo	570 & 153
traU	993	$traU^{B}$	969	$traU^{B}$	981
trbC	639	trbC	435	trbC	468
traN	1809	traN	1812	traN	1827
traF	774	traF	903	traF	912
traQ	285	not determined		not determined	
traH	1377	$traH^{E}$	1341	$traH^{E}$	1341
traG	2817	$traG^{B}$	2718	$traG^{B}$	2730
traS	522	not determined		not determined	
traT	735	not determined		not determined	
traD	2154	$traD_F^{E}$	1689	$traD_F^{E}$	1707
traI	5271	$traA_{Ti}^{BE}$*	4152	$traA_{Ti}$	4137
$traX^{E}$	747	not determined		not determined	
not determine		$traD_{Ti}$	294	$traD_{Ti}$	297

Operons were predicted by DOOR: Database for prOkaryotic OpeRons. From top to bottom of the list of tra genes, the beginning of a predicted operon is indicated by "B" followed by the end of the same operon as indicated by "E".* = predicted operon begin on $traA_{Ti}$ and end on RBE_0422 (putative transposase).

Chapter 3: Identifying the best reference gene for transcriptional analysis of the *tra* genes of *Rickettsia bellii* RML 369-C

Introduction

Rickettsiae are obligate-intracellular, coccobacillary, Gram-negative bacteria that include invertebrate symbionts. Genetic analysis of rickettsiae has revealed that plasmids are widely distributed within the genus (Ogata *et al.* 2005a; Baldridge *et al.* 2010) and *tra* genes are common (Weinert, Welch and Jiggins 2009), suggesting that rickettsiae can transfer mobile genetic elements, i.e. plasmids, to receptive bacteria via conjugation. *Rickettsia bellii* strain RML 369-C (Ogata *et al.* 2006), *Rickettsia massiliae* MTU5 (Blanc *et al.* 2007b), *Rickettsia felis* strain LSU-lb (Gillespie *et al.* 2014), and the rickettsial endosymbiont of *Ixodes scapularis* (REIS, i.e. *Rickettsia buchneri*) (Gillespie *et al.* 2012; Kurtti *et al.* 2015) possess a set of *tra* genes that encodes an apparently complete bacterial conjugation system, but it is not known whether they are functional. In contrast, several other rickettsiae, e.g. *Rickettsia felis* strain California 2, only encode a few of the *tra* genes, suggestive of an incomplete conjugation system (Ogata *et al.* 2005a). Nevertheless, pili-like structures have been observed in *R. felis* as well as *R. bellii* (Ogata *et al.* 2005a, 2005b, 2006). Furthermore, the *tra* clusters of *R. bellii* RML 369-C, *R. felis* LSU-lb, and *R. massiliae* MTU are similar at the nucleotide sequence level, gene order, and gene orientation to at least one of the *tra* clusters of *R. buchneri*, suggesting multiple lateral gene transfer events between rickettsiae (Blanc *et al.* 2007a; Gillespie *et al.* 2012). Despite this, bacterial conjugation has not been documented in any rickettsiae.

The initiation of conjugation in *Escherichia coli*, the model organism, begins with pilus contact with a recipient cell followed by retraction of the pilus and the formation of

a stable mating pair (Firth and Skurray 1992). The timing and signal to process the transfer DNA is unknown; nevertheless, the mechanism of generating the transfer DNA has been established. Briefly, TraI is the main processing enzyme with nickase and helicase activity (Matson, Sampson and Byrd 2001; Matson and Ragonese 2005) that covalently binds to the transfer DNA with the help of accessory proteins (Howard, Nelson and Matson 1995; Nelson *et al.* 1995; Ragonese *et al.* 2007). Subsequently, TraI-bound DNA docks with TraD (Lang *et al.* 2011), an inner membrane coupling protein that initiates the helicase activity of TraI to generate a single-stranded transfer DNA. TraD then opens the pore for transport of the nucleoprotein complex to the recipient thus completing conjugation. *E. coli* has been shown to express TraI until late stationary phase; however, mating success was highest from early lag phase to mid exponential phase (Frost and Manchak 1998). TraI shows homology to $TraA_{Ti}$ in *R. bellii* at the protein level, sharing similar domains with presumably similar functions (Fig. 3.1) as well as belonging to the same class of enzyme, relaxases (Francia *et al.* 2004). As seen in *E. coli* (Frost and Manchak 1998), certain growth phases of rickettsiae may correlate with transcription of $traA_{Ti}$ and other *tra* genes, suggesting that conjugation occurs around the time of highest transcription.

The aim of this research was to follow the transcriptional patterns of $traA_{Ti}$ in *R. bellii* in comparison with stably transcribed reference gene(s) to facilitate the assessment of the function of the *tra* genes. We determined the best reference gene at different times of growth in ISE6 (*I. scapularis* tick embryonic cells), Vero (African green monkey kidney cells) and L929 (mouse fibroblast cells), all of which support continuous *R. bellii* replication, and may be taken to represent the rickettsial life cycle alternating between the

arthropod and the mammalian host. A two-step quantitative reverse transcription PCR (qRT-PCR) and two specialized statistical programs, Normfinder (Andersen *et al*., 2004) and BestKeeper (Pfaffl *et al*., 2004), were used to screen 10 housekeeping genes in *R. bellii*: *16S rRNA, atpB, dnaK, gltA, gyrA, infB, metG, nrdF, rpoB,* and *tlc5* (Table 3.1). Of these, *metG* was found to be the most stably transcribed throughout the infection cycles. We validated this as our reference gene by using it to demonstrate the differential transcription patterns of *rickA* and *sca2* as reported (Reed *et al*., 2014). We then used *metG* to analyze the transcription of the components of *R. bellii tra* cluster, $traA_{Ti}$, $traD_{Ti}$, *traL, traE, traB, traV, traC, traW, traU, trbC, traN, traF, traH, traG,* $traD_{F}$, and RBE_0422 during growth in ISE6 cells, and we used *metG*, *nrdF*, and a combination of *metG* and *nrdF* for a comparative analysis of transcription activity of $traA_{Ti}$ during replication in different host cells in vitro.

Materials and Methods

Determining the domains of $TraA_{Ti}$ and related proteins

The location and functions of the domains of $TraA_{Ti}$ of *R. bellii* RML 369-C (YP_537591.1), TraI of *R. buchneri* (KDO03561.1), TraA of *Agrobacterium tumefaciens* (WP_035228570.1), and TraI of *E. coli* (WP_038999217.1) were identified using Batch CDD search (Marchler-Bauer and Bryant 2004) on the Conserved Domain Database site (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). The graphical representations of the domains of each protein were drawn in Adobe Photoshop to illustrate the location and overall function of each segment of the proteins (Fig. 3.1).

Host cells

Cell line ISE6 (Munderloh *et al.* 1999) was cultured at 34^{o}C using a modified Leibovitz's L15 medium (L15C-300) supplemented with 5% fetal bovine serum (FBS), 5% tryptose phosphate broth, and 0.1% lipoprotein concentrate as described (Munderloh and Kurtti 1989; Oliver *et al.* 2013). Vero (ATCC CCL-81) and L929 (ATCC CCL-1) were cultured at 37^{o}C using RPMI+GlutaMAX 1640 medium (Gibco, Grand Island, NY) supplemented with 10% FBS and 2 mM L-glutamine.

Preparation of cell-free *R. bellii* stabilate for inoculation of host cell cultures

R. bellii RML 369-C was grown in ISE6, Vero and L929 cells in 25 cm^2 tissue culture flasks (BD Falcon, NJ) at 34^{o}C in L15-C300 supplemented with 10% FBS for at least 2 subcultures within the respective host cells before preparing and freezing cell-free rickettsia stocks. In brief, infected cells were scraped off the flask, suspended in L15C-300 medium, transferred to 1.5 ml microcentrifuge tubes containing 60-90 grit silicone carbide (Lortone, Inc., Mukilteo, WA), and vortexed at maximum speed for at least 20 seconds to rupture host cells and obtain cell-free *R. bellii*. Liberated *R. bellii* were filter-purified using a Whatman 2.0 µm filter (GE Healthcare Life Science, NJ) as described in Oliver *et al.* (2014), collected by centrifugation at 13,000x rcf for 5 minutes at 4^{o}C and resuspended in freezing medium (L15C-300 with 50% FBS and 10% DMSO), then aliquoted into cryogenic vials and frozen at a controlled rate of -1^{o}C/min for storage in liquid nitrogen.

Inoculation of host cell cultures

Uninfected cells were seeded into 12.5 cm^2 tissue culture flasks (BD Falcon, NJ) with 3 flasks for each time point and cells were allowed to attach overnight. Before inoculation, the multiplicity of infection was determined by counting rickettsiae and host

cells. Rickettsiae were visualized using LIVE/DEAD BacLight Viability Kit (Molecular Probes, OR) and quantified using a Petroff-Hauser counting chamber. The number of host cells in triplicate cultures was quantified using an Improved Neubauer counting chamber. The number of live rickettsia divided by the host cell counts yielded the multiplicity of infection. Subsequently, the rickettsial stabilates were diluted to the point where inoculating host cell layers with a total volume of 200 µl resulted in 10-50 rickettsiae per host cell. Immediately following inoculation, cultures were transferred to 4^{o}C for 1 hour to allow binding of rickettsiae to cells and to synchronize the infection. Cell cultures were shifted from 4^{o}C to 34^{o}C and cultures incubated at 34^{o}C for 1 hour to facilitate endocytosis of bound rickettsiae. Cell layers were washed twice with 1x phosphate buffered saline containing calcium and magnesium to remove unbound rickettsiae. Fresh L15C-300 with 10% FBS was added to infected cell layers and cultures incubated at 34^{o}C for 12, 24, 36, 48, and 72 HPI for infected ISE6 cultures and 12, 24, 36, 48, 60, 72, 84, and 96 HPI for infected Vero and L929 cultures.

Quantification of *R. bellii*

Quantification methods were validated by comparing purified cell-free *R. bellii* (rickettsiae prepared, RP) with yields from host cells infected with *R. bellii* (whole cell, WC) at 12 and 72 HPI. The total number of *R. bellii* obtained from each extraction method was determined by qPCR, using a standard curve based on quantification of a single copy gene, *gltA*. Statistical significance was measured using a two-sided two sample Student's t-test with significance set at an alpha of 0.05.

Rickettsial DNA and RNA were extracted from each culture. Whole cell lysis of rickettsiae and host cells was initially used for preparing total DNA and RNA. However,

first strand cDNA synthesis using random hexamer yielded low rickettsial cDNA at earlier times post inoculation of rickettsiae; thus, cell-free rickettsial preparation were used. Cell cultures were harvested and RP were prepared with silicone grit as stated previously and stored at -80^{o}C. Pellets were thawed and immediately lysed for DNA and RNA extraction using ZR-Duet DNA/RNA MiniPrep kit (Zymo Research, CA) according to the manufacturer's protocols. Extracted DNA and RNA were stored at -20^{o}C and -80^{o}C, respectively. qPCR was used to quantify the *gltA* copy numbers from purified DNA of liberated *R. bellii* grown in ISE6, Vero, and L929 (Eremeeva, Dasch and Silverman 2003; Baldridge *et al.* 2010). The pCR^{TM}4 TOPO$^{®}$ (Invitrogen, NY) plasmid fused with the PCR product of Rbellii qCSF/R primers (Table 3.1) and the respective primers were used to generate the standard curve for quantifying absolute copy numbers of *R. bellii* for each time point using Brilliant II SYBR$^{®}$ Green qPCR Low Rox Master Mix (Agilent, CA) at 1x with a final concentration of 0.08 µM of each primer, and 5 µl of the total eluted *R. bellii* DNA. The cycling parameters were 95^{o}C for 10 minutes, followed by 40 cycles of 95^{o}C for 30 seconds, 55^{o}C for 1 minute, and 72^{o}C for 30 seconds. A dissociation curve was used to confirm a single PCR product with a cycle starting at 95^{o}C for 1 minute, 55^{o}C for 30 seconds, and 95^{o}C for 30 seconds immediately after the quantification cycles.

Growth of *R. bellii* in tick and mammalian cell cultures

Cell cultures infected with *R. bellii* were extracted for DNA at 12, 24, 36, 48, and 72 HPI for ISE6 and additional times of 60, 84, and 96 HPI for Vero and L929. Rickettsiae numbers were quantified as previously mentioned. The values from triplicate

cultures at each time point were averaged and graphed using SigmaPlot 13 (Systat Software, Inc., CA).

Transcript levels of reference genes extrapolated from tiling microarray analysis of *Rickettsia rickettsii*

R. bellii reference genes (*metG*, *nrdF*, *gyrA*, *gltA*, and *16s rRNA* gene) were inferred to be transcribed above background by extrapolation from tiling array data of RrSS orthologous genes. Triplicate cultures of Vero cells in 75 cm^2-flasks were grown to confluency at 37°C in RPMI1640 medium supplemented with 10% FBS, 25 mM HEPES buffer, and 0.25% $NaHCO_3$, pH 7.5, inoculated with *Rickettsia rickettsii* strain Sheila Smith (RrSS) and incubated until cell layers were completely infected (3-4 days at the University of Minnesota Biosafety Level 3 Research Laboratory). RNA was then purified from these and from uninfected control cultures using Tri-Reagent (Sigma #T9424) according to the product protocol, briefly as follows: cells were dissolved in TriReagent; the aqueous phase was combined with an equal volume of isopropanol; the precipitated RNA was washed in 75% ethanol and then resuspended in water. RNA was directly labeled using the Kreatech ULS Labeling Kit (Leica Biosystem) according to the kit protocol. This method labels guanine nucleotides in RNA using platinum complexes linked with Cy3 fluorophores. Agilent (Santa Clara, CA) manufactured the whole genome tiling array for RrSS [GenBank CP000848.1] with designs we made using eArray, an array design tool provided by Agilent. Each slide consisted of eight arrays with 60,000 60mer oligonucleotides (probes), each with 15 nucleotides of overlap. The triplicate array data were quantile normalized and T-tested to determine significance (0.05). We used Artemis (http://www.sanger.ac.uk/) to help analyze the results of the

arrays by superimposing the mean quantile normalized signal levels from each of the hybridized probes onto a graphical representation of the RrSS genome, such that transcription levels for each of the overlapping probes could be viewed in an entire genome context. The level of transcription is represented by the sum of the intensity of each probe within an annotated gene, using the calculated "area under the curve". Numerical values of the "area under the curve" are displayed in Table 3.2.

Selection, quantification, and validation of *R. bellii* reference gene transcripts

To identify stable reference genes to analyze the transcriptional responses of the *tra* genes in *R. bellii* grown for different times (12, 24, 36, 48, and 72 HPI) and in different host cells (48 HPI only), a two-step qRT-PCR was used employing sample maximization (Hellemans *et al.* 2007). Reference genes (*16S rRNA*, *atpB*, *dnaK*, *gltA*, *gyrA*, *infB*, *metG*, *nrdF*, *rpoB*, and *tlc5*) were selected based on their predicted molecular function in different metabolic classes and their property as housekeeping genes (Table 3.1). In addition, five out of the ten reference genes were selected to show transcription levels above background in *Rickettsia rickettsii* Sheila Smith based on tiling array data (Fig. 3.3, Table 3.2).

First strand cDNA was synthesized from total RNA using random hexamer primers (Integrated DNA Technology, IA) and Superscript II Reverse Transcriptase (Life Technologies, NY) following the manufacturer's protocol and stored at -20^{o}C. All cDNA was diluted 1:20 with water and 5 µl of the dilution was added to each qPCR reaction. Master mix and cycling parameters were similar to the conditions used to determine *gltA* copy numbers, except the primer hybridization temperature for *traV* detection was lowered to 50^{o}C based on the GC content of the primer set used. The transcript numbers

were normalized using the ratio of *gltA* genomic copy numbers detected at all time points to 12 HPI.

Normfinder and BestKeeper (http://moma.dk/normfinder-software and http://www.gene-quantification.com/bestkeeper.html, respectively) were used to rank and identify the best reference gene or combination of reference genes to use for comparative analysis. Normfinder uses a model based approach to determine the most stable gene (Andersen *et al.* 2004) and can handle high variations from low copy numbers of transcript from earlier time points because of lower number of rickettsiae and high copy numbers at later times. The ranking of our genes using BestKeeper was based on the highest coefficient of correlation for the best reference gene to the lowest coefficient of correlation for the least favorable reference gene (Pfaffl *et al.* 2004). BPROM, a promoter prediction program (http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb), was used to identify potentially shared promoters upstream of the best reference genes identified by BestKeeper to discount similarly regulated genes for use as reference genes.

Transcription of *rickA* and *sca2* in Vero cells to validate *metG* as a reference gene

Validation of *metG* as the best reference gene was accomplished using primers to determine the transcription pattern of *rickA* and *sca2*. The differences within and between the transcription patterns of *rickA* and *sca2* were statistically analyzed using multiple comparisons with an ANOVA test and a Holm-Sidak correction with an alpha value set at 0.05 in SigmaPlot 13.

Relative transcription of $traA_{Ti}$ to reference genes

First strand cDNA synthesis was generated, as mentioned previously, from RNA of *R. bellii* isolated from infected ISE6, Vero, and L929 cells at all the time points. Two-step qRT-PCR was done using primers for $traA_{Ti}$ and for the reference genes *metG* and a combination of *metG* and *nrdF*. Template consisted of 1st strand cDNA synthesized at all time points in all cell lines. The $2^{-\Delta\Delta CT}$ method was used to quantify fold changes of $traA_{Ti}$*:metG* at all time points relative to 12 HPI (Livak and Schmittgen 2001). SigmaPlot 13 was used to generate line graphs.

Relative transcription of *metG* and *nrdF* using $traA_{Ti}$ as a reference gene

The C_T values from *metG*, *nrdF*, and $traA_{Ti}$ from *R. bellii* grown in ISE6 were used **only** to demonstrate the constitutive transcription of *metG* and *nrdF* relative to $traA_{Ti}$ and **not** to imply that $traA_{Ti}$ is a good reference gene. The graph was created using SigmaPlot 13.

Transcriptional analysis of the *tra* from 12 to 72 HPI

Transcriptional analysis of $traA_{Ti}$, $traD_{Ti}$, *traL*, *traE*, *traB*, *traV*, *traC*, *traW*, *trbC*, *traU*, *traF*, *traH*, *traG*, $traD_F$, and RBE_0422 was carried out under similar conditions and cycling parameters, as previously stated, using cDNA from 12 and 72 HPI of *R. bellii* grown in ISE6 as templates together with respective primer sets (Table 3.1).

Results

Domains of *R. bellii* $TraA_{Ti}$ are related to conjugative relaxases of other bacterial species

The domains of *R. bellii* $TraA_{Ti}$ are depicted in Fig. 3.1. The MobA/L, ABC ATPase, and UvrD C domains of *R. bellii* $TraA_{Ti}$ and *R. buchneri* TraI constitute the conjugal segment of TraA. This conjugal segment is predicted to have similar function as

the conjugal segment of *A. tumefaciens* TraA and *E. coli* TraI (underlined in Fig. 3.1) such as DNA binding, nicking, and unwinding (Alt-Morbe *et al.* 1996; Matson, Sampson and Byrd 2001; Matson and Ragonese 2005) when generating an ssDNA for DNA transfer, which are all qualities of relaxases although the amino acid sequences may vary between the families of relaxases (Garcillan-Barcia, Francia and de la Cruz 2009) In addition to the conjugal segment of $TraA_{Ti}$ of *R. bellii*, a primase was identified with a Toprim domain at the carboxyl terminal end. The overall domains of *R. bellii* $TraA_{Ti}$ are similar to the well characterized conjugal relaxases of other bacteria; thus, we suspect that *R. bellii* $TraA_{Ti}$ expression may be a good indicator of the time of induction for potential bacterial conjugation.

Growth of *R. bellii* in tick and mammalian cell cultures

To verify that there were no differences in rickettsial numbers determined from RP vs. WC, *R. bellii* DNA was prepared and the *gltA* genomic copy numbers in each preparation were compared (Table 3.3). According to the two-sided two sample Student's t-test with alpha set to 0.05, no significant differences were observed between the two isolation procedures. Thus, the process of preparing cell-free rickettsial preparations did not cause a significant loss of rickettsiae when compared with whole cell lysate.

A two-sided two sample Student's t-test with an alpha value at 0.05 was used to compare the number of detected *gltA* copy numbers in rickettsiae prepared (RP) by semi-purification and whole cell (WC) lysis of *R. bellii* RML 369-C grown in ISE6 at 34^{o}C. Results show no significant difference between the two methods.

Subsequently, *gltA* copy numbers were quantified using the RP method and qPCR for all time points and host cells. The growth curves for *R. bellii* in different cell lines

(Fig. 3.2) shared similar growth phases and a doubling time of approximately 8 hours during the period of 36 to 60 HPI similar to times reported for *Rickettsia prowazekii* and *Rickettsia rickettsii* (Hopps *et al.* 1959; Wisseman Jr and Waddell 1975; Wisseman Jr *et al.* 1976). The initial higher copy numbers of *gltA* from *R. bellii* isolated from L929 could have resulted from inefficient washing or highly efficient phagocytosis and killing of rickettsiae, but slower digestion of rickettsial DNA by L929. We concluded that the replication of *R. bellii* was similar in the different host cells using an MOI of 10-50 followed by a wash step.

Identifying the best reference gene and best combination of genes

Ten *R. bellii* housekeeping genes were selected based on their different functional classes, i.e. translation, DNA replication: *16S rRNA*, *atpB*, *dnaA*, *gltA*, *gyrA*, *infB*, *metG*, *nrdF*, *rpoB*, and *tlc5* (Table 3.1). Results for *R. rickettsii* microarray analysis showed transcription levels above background for each of the selected genes (Fig. 3.3). Normfinder and BestKeeper were used to analyze the transcript data of the *R. bellii* reference genes (Table 3.4). Normfinder ranked *metG* as the most stably transcribed gene throughout a 72 HPI growth period in ISE6 and at 48 HPI between the different host cell lines. Normfinder identified a combination of *metG* and *nrdF* for transcriptional analysis throughout a 72 HPI in ISE6, while *gyrA* and *metG* were selected by Normfinder when comparing between host cells at 48 HPI. In comparison, BestKeeper ranked *gyrA* and *metG* as the top two genes to use for a time point analysis and for comparing between host cells based on the coefficient of correlation (Table 3.4). A caveat of BestKeeper is that it can identify reference genes that are transcriptionally co-regulated, which is not ideal when identifying the best reference gene. To avoid co-regulated transcription, the

selected genes were chosen based on their different molecular functions that belong to different metabolic classes, i.e. *gyrA* and *metG* functions in DNA replication and initiation of translation, respectively, suggesting that the genes tested are not co-regulated. Analysis of the genetic sequences upstream of *metG* and *gyrA* using BPROM identified a promoter upstream of *gyrA* at position 879453 with a linear discriminant function score of 9.77, and located a -10 box at 879468 with a score of 61, and a -35 box at 879492. Also, *metG* has a promoter at position 282683 with linear discriminant function score of 9.18, a -10 box at 282668 with score of 63, and a -35 box at 282649 with a score of 47. Both promoters were shown to have several transcription factor binding sites, but share only a leucine-responsive regulatory protein binding site. This transcription factor is involved in controlling the cell response to nutritional stress, particularly to leucine levels (Cho *et al.* 2008). However, *R. bellii* does not encode a leucine-responsive regulatory protein-like protein according to the annotated *R. bellii* RML 369-C genome, indicating that *metG* and *gyrA* are not similarly regulated. Overall, *metG* is the best reference gene selected by Normfinder and BestKeeper either as the best or second best gene to use as a reference when comparing rickettsiae replicating in different host cells.

Transcription of *rickA* and *sca2* is induced at different growth phases in Vero cells, initially validating *metG* as a reference gene

R. bellii uses actin-based motility similar to other intracellular bacteria (Oliver *et al.* 2013). Thus far, *rickA* and *sca2* were found to be involved in actin tail polymerization (Gouin *et al.* 2004; Haglund *et al.* 2010), and their products, RickA and Sca2, were shown to be expressed during distinct phases of intracellular growth (Reed *et al.* 2014).

Therefore, we were interested in defining the transcriptional patterns of *rickA* and *sca2* to validate our reference gene. Relative transcription of *rickA* and *sca2* was determined as described using *metG* as the reference gene. Up-regulation of *sca2* was observed from 12 to 24 HPI (Fig. 3.4A), which correlated with the lag phase of the growth curve (Fig. 3.2). *rickA* showed a trend of up-regulation between 48 to 96 HPI with significant differences in transcription levels at 72 and 96 HPI that coincided with the mid-log phase to the stationary phase. *sca2* showed a significant difference in relative transcription level between 12 and 24 HPI that coincided with the lag phase and between 24 and 72 HPI that coincided with the beginning of the log phase and stationary phase, respectively (Fig. 3.4A). The transcription patterns of *sca2* and *rickA* were modulated by distinct phases of infection and growth. Analyses of *rickA* and *sca2* transcription showed significant differences between each other at 24 HPI (Fig. 3.4B). Furthermore, *rickA* and *sca2* transcript levels displayed a trend similar to the growth curve (Fig. 3.1 and 3.4B). Thus, *sca2* may play an important role during the lag phase of invasion especially during the beginning of the log phase, while *rickA* may have an important role in the exponential phase of *R. bellii* replication.

Transcription of $traA_{Ti}$ in *R. bellii* is induced in ISE6 at 72 HPI

Two-step qRT-PCR was used to quantify $traA_{Ti}$ mRNA through time in different host cells using *metG* and a combination of *metG* and *nrdF* as reference genes. At 24, 36, and 48 HPI $traA_{Ti}$ in *R. bellii* grown in ISE6 did not show any apparent differences in transcription when compared to 12 HPI using *metG* only, *nrdF* only, or *metG* and *nrdF* as reference genes, but at 72 HPI $traA_{Ti}$ was up-regulated compared to 12 HPI (Fig. 3.5). In addition, $traA_{Ti}$ was also up-regulated in *R. bellii* grown in ISE6 in comparison to *R. bellii*

grown in Vero and L929 at 72 HPI (Fig. 3.5). As a result of the induction of $traA_{Ti}$ observed at 72 HPI, *R. bellii* was grown in ISE6 for 12 and 72 HPI and examined for the transcriptional patterns of the other annotated *tra* genes. All examined *tra* genes were found to be up-regulated at 72 HPI (Table 3.5) as indicated by the lower C_T values where the threshold is reached with lower number of cycles when starting with a greater quantity of template.

Discussion

Several studies have analyzed relative transcription patterns in rickettsiae, e.g. using *16S rDNA*, *metG*, and *gltA* (Gaywee *et al.* 2002; Galletti *et al.* 2013; Oliver *et al.* 2013). However, the reference genes chosen were not rigorously evaluated and validated as the most stably transcribed genes under experimental conditions using statistical programs. In this study, we evaluated 10 different genes and identified and validated *metG* or *metG* and *nrdF* as the best reference genes to use for transcriptional analysis of the *tra* genes during *R. bellii* growth in ISE6 cell cultures and in different mammalian cell lines using NormFinder. BestKeeper identified *gyrA* and *metG* as the top two reference genes. Based on the promoters (predicted by BPROM) and the differences in molecular functions between *metG* and *gyrA*, we surmised that they were unlikely to be transcriptionally co-regulated. *metG* and *nrdF* were also shown to be constitutively transcribed in *R. bellii* (Fig. 3.6). Tiling array data from *R. rickettsii* also showed that transcription levels of *metG*, *nrdF*, *gyrA*, *gltA*, and *16S rDNA* were above background (Fig. 3.3). Both results corroborated the stability of the transcription of *metG*, *nrdF*, and *gyrA* indicating that all three may be good reference genes. However, *metG* consistently ranked as the best reference gene to be used for relative transcriptional analysis. *metG*

was also used to determine the transcriptional response of *R. rickettsii* to a shift in temperature suggesting that both pathogenic and nonpathogenic rickettsiae may transcribe *metG* at stable levels under different treatment conditions.

We further validated our reference genes by tracking transcriptional patterns of *sca2* and *rickA*. We observed a differential transcription pattern of *sca2* and *rickA*, with a distinctive up-regulation of *sca2* during the lag phase of rickettsial growth and a trend of increasing variance of the transcription of *rickA* during exponential phase. Sca2 up-regulation at 24 HPI suggested that it played an important role when *R. bellii* was invading and adapting to the host cells until it was ready for the exponential phase (Fig. 3.4A). In contrast, *rickA* demonstrated high variability of transcription levels during the mid-exponential growth phase to the stationary phase suggesting discontinuous and variable motility of *R. bellii* (Fig. 3.4A). This variability of *rickA* transcript may be attributed to the first asynchronous infection of neighboring cells near 48 HPI when rickettsiae begin to spread and infect new cells (Wisseman Jr and Waddell 1975; Silverman, Wisseman Jr and Waddell 1980). Both *rickA* and *sca2* had general transcription patterns that trended with the growth curve of *R. bellii* except at the beginning of the exponential phase (24 HPI) where *sca2* transcription was significantly different from *rickA* (Fig. 3.4B). Sca2 is found at the surface (Sears *et al.* 2012) and has been shown to be involved in adhesion to mammalian cells (Cardwell and Martinez 2009). Thus, *sca2* is probably transcribed significantly more than *rickA* at the initial part of the exponential phase to express enough Sca2 to ensure that they can continue to adhere to and invade host cells, as well as to enable them to use actin-base motility for later stages of cell-to-cell spread as indicated by Reed *et al*. 2014. Expression patterns of

RickA and Sca2 in *Rickettsia parkeri* were significantly different at early and late infections correlating with actin-tail morphology and were observed along with *rickA* or *sca2* knockouts to determine the differences in cell-to-cell spread (Reed *et al.* 2014). In contrast, our results indicated that *rickA* and *sca2* were differentially transcribed at different times post infection highlighting the importance of *rickA* and *sca2* at different growth phases of *R. bellii*. Furthermore, Sca2 of *R. bellii* lacks the formin homology 2 domain, but maintains the formin homology 1 domain, and has been shown to sufficiently accelerate actin polymerization indicating it functions as a tandem monomer-binding nucleator (Haglund 2011). These differences may result in different regulation of expression and function of *sca2* and indirectly affecting *rickA*. Nonetheless, we demonstrated that a significant difference in the transcription levels of *rickA* and *sca2* of *R. bellii* existed during the transition of lag to exponential growth phase. Thus, this is the initial validation of *metG* as a suitable reference gene.

$TraA_{Ti}$ was chosen as the indicator of the conditions to examine transcription of the other *tra* genes because it was shown to reflect responses of the entire *tra* cluster. $TraA_{Ti}$ has domains that suggest it is functionally similar to TraI (Fig. 3.1), indicative of its role in transfer DNA synthesis (Matson, Sampson and Byrd 2001; Matson and Ragonese 2005). Specifically, $TraA_{Ti}$ of *R. bellii* has a MobA/L domain that probably has nickase activity similarly to the TrwC domain in *E. coli*, and the UrvD domain is shared by both *R. bellii* and *E. coli*. The main difference is that *R. bellii* $TraA_{Ti}$ has a primase domain not present in TraI of *E. coli* (Fig. 3.1). A possible role of primase in *R. bellii* $TraA_{Ti}$ may be the initiation of DNA replication immediately following the generation of transfer ssDNA in the donor, or initiation of replication of the transfer DNA in the

recipient. Thus far, the primase domain is unique to the relaxases of rickettsiae and insight into the function of this domain may identify a different mechanism in conjugation.

The conditions that stimulated high transcriptional activity of $traA_{Ti}$ using *metG* as the reference gene were met when *R. bellii* was growing in ISE6 at 34^{o}C for 72 hours (Fig. 3.5), near the beginning of the stationary phase (Fig. 3.1). In contrast, *R. bellii* did not show any changes in $traA_{Ti}$ transcriptional activity when grown in mammalian cells. Possibly, cellular and metabolic processes of mammalian cells inhibited or indirectly altered the transcription of $traA_{Ti}$. In nature, *R. bellii* has been observed to primarily infect arthropods, particularly ticks, so the increase in relative transcription found when grown in a tick cell line in the early stationary phase (Fig. 3.2, 72 HPI) suggested the importance of the tick host in possible bacterial conjugation. Timing of transcription during the early stationary phase may prepare the bacteria for conjugal activity during this phase. This is in contrast to *E. coli* in which mating efficiency is maximal during rapid replication (mid-log phase) when *tra* transcripts are still detected although Tra proteins that are components of the transfer structure remain stable into stationary phase when TraM and TraI are lost (Frost and Manchak 1998).

The transcription of the other annotated *tra* genes analyzed was positively correlated with $traA_{Ti}$. Although $traD_{F}$ showed lower transcription activity overall, it still correlated positively with the other *tra* genes (Table 3.5), while transcription of RBE_0422 showed no correlation with the other *tra* genes despite the prediction that RBE_0422 and $traA_{Ti}$ were part of an operon (Table 2.3, (Dam *et al.* 2007; Mao *et al.* 2009, 2014)). All the *tra* genes were simultaneously up-regulated at early stationary

phase or 72 HPI, further validating *metG* as the best reference gene (Table 3.5). *traF* mRNA was not detectable during the lag phase or at 12 HPI, but was present at the early stationary phase indicating that *traF* may not be important during the lag phase, but plays a role near the early stationary phase. Also, this shows that not all *tra* genes were transcribed together and may have different transcriptional regulation. A mutation in *traF* in *E. coli* eliminated formation of the pilus but did not affect replication, suggesting that pilus formation is not important for growth in *E. coli* (Anthony *et al.* 1999), and this may also be true for *R. bellii traF*. *Rickettsia bellii traG* showed highest relative transcription near the early stationary phase indicating the importance of *traG* at this time and life stage. In *E. coli*, *traG* is multifunctional and participates in pilus assembly, mating pair stabilization, and entry exclusion, which further highlights this time point and stage of growth as a potential period to target for testing conjugation in *R. bellii* (Manning, Morelli and Achtman 1981; Firth and Skurray 1992; Marrero and Waldor 2005; Audette *et al.* 2007). Taken together, our results suggest that testing for conjugation in *R. bellii* should be done in tick cell culture at later stages of infection.

Determining if the *tra* genes are functional as operons may be useful for understanding the regulation of these genes. There was evidence indicating that *traB*, -*V*, and -*C* may form an operon as suggested by overlapping of the start and stop codon of adjacent genes and similar degree of transcriptional up-regulation. Further evidence from the Database for prOkaryotic OperRons support *traB*, -*V*, and -*C* as an operon that also includes *traL* and -*E* (Table 2.3). The Database for prOkaryotic OpeRons identified another *tra* operon that included *traU*, *trbC*, *traN*, -*F*, and -*H* (Table 2.3) and genetic evidence show overlapping start and stop codons from *trbC*, *traN*, -*F*, and -*H* supporting

these *tra* genes as belonging to an operon. The C_T values and fold increase of *traU*, *trbC*, *traN*, *-F*, and *-H* differed from each other despite evidence of an operon. However, differential and "segregational" decay of polycistronic transcripts are common and may explain the transcriptional pattern observed. The stability of transcripts can be affected by the secondary structure of RNA and/or the frequency of translation (Koraimann *et al.* 1996; Kristoffersen *et al.* 2012) highlighting the importance of proteomic analysis to support the regulation of transcription.

We considered RBE_0435 of *R. bellii* RML 369-C (Fig. 3.7) to be a possible candidate for the putative gene encoding the propilin (or pilin) subunits of the pilus. Searching for this protein homolog is difficult considering this coding gene is poorly conserved (Lawley *et al.* 2003) although key characters such as a signal peptide and two transmembrane-spanning domain can help to identify potential homologs as seen with the homologous protein VirB2 (Fig. 3.7, (Gillespie *et al.* 2009)). Interestingly, the proteins involved in directing pilin to the periplasm (TraQ in *E. coli*) and the enzyme responsible for N-acetylation of pilin (TraX in *E. coli*) have not been found in *R. bellii* (Table 2.3). This suggests that this potential pilin protein either does not require the modification and/or that it may not be involved in forming pili. Alternatively, the pili-like structures that have been observed in *R. bellii* (Ogata *et al.* 2006) could originate from a different system, i.e. the VirB/D system. Contrarily, RBE_0435 locus has been illustrated to be the carboxyl-terminal end of *traK* (Gillespie *et al.* 2012). The pili aspect of the *tra* system remains elusive.

If rickettsiae can engage in lateral gene transfer, then this could contribute new genes to the rickettsial genome, buffer the rate of reductive genome evolution (Blanc *et*

al. 2007b) and enhance rickettsial fitness in new environments. The *tra* genes can facilitate import of new genes through recombination and insertion at a particular hot spot between $traA_{Ti}$ and $traD_{F}$, as suggested by Blanc et al. (Blanc *et al.* 2007a). Evidence of this phenomenon is illustrated by the *tra* genes of *R. buchneri* where genes typical of Gram-positive aminoglycoside antibiotic biosynthesis gene clusters were found in a similar hot spot suggesting that the *tra* cluster can support piggybacking genes into the rickettsial genome to increase fitness and slow down reductive genome evolution (Gillespie *et al.* 2012). Weinert *et al.* (2009) have pointed out that there is no evidence for the preferential transfer of genes between closely related species of rickettsiae and speculate that instead they may engage in "regenerative horizontal" DNA exchange further combatting a reductive genome evolution. This could be achieved by recombination of foreign DNA into the rickettsial genome, but recombination events have been argued to be generally rare in *Rickettsia* (Weinert *et al.* 2009), although they seem more common in some species, e.g. *R. bellii* (Hernandez-Lopez *et al.* 2013). The possibility that the *tra* gene cluster can mediate acquisition of new genes and act as a hot spot for insertion to assist in the repair of damaged DNA in *R. bellii* through natural transformation is intriguing.

There is, however, ample evidence that some rickettsiae, especially those symbiotically associated with arthropods, harbor horizontally acquired genetic material on their plasmids that may have been acquired via bacterial conjugation. For example, *Rickettsia peacockii* maintains five genes in the pRPR plasmid that are distantly related to the *Pseudomonas aeruginosa* glycosylation island and may function in the uptake of dihydroxyacetone phosphate for phospholipid biosynthesis (Felsheim, Kurtti and

Munderloh 2009). Furthermore, *R. buchneri* encodes a complete biotin operon on pREIS2 that shares homology with biotin operons recently acquired by horizontal gene transfer in *Neorickettsia risticii*, *N. sennetsu* and *Lawsonia intracellularis* (Gillespie *et al.* 2012). Recently, an RTX type I secretion system (TISS) was found in *R. felis* plasmid pLSU-lb that was closely related to the TISS genes of *Cardinium* endosymbiont cBtQ1 of *Bemisia tabaci*, in which the RTX TISS genes were thought to originate through lateral gene transfer from *Vibrio* species (Gillespie *et al.* 2014). In addition, other *R. bellii* isolates contain plasmids that have yet to be sequenced (Baldridge *et al.* 2010), and that may contribute to the dissemination of operons that mediate symbiosis. There are other mobile genetic elements such as the integrated genetic element found on the *R. buchneri* chromosome that was supported by bioinformatics to be closely related to and include the conjugation genes of *R. bellii* (Gillespie *et al.* 2012). The acquisition of the *tra* genes may have been via a plasmid based on the recent discovery that *R. felis* LSU-lb harbors a plasmid containing the *tra* cluster closely related to pREIS3 of *R. buchneri*. The mechanism of transfer of an integrative conjugative element requires a functional integrase or recombinase that will excise the sequence from the chromosome followed by transfer strand synthesis and transport into the recipient either by encoding its own conjugation system or parasitizing another present in the donor (Wozniak and Waldor 2010; Bellanger *et al.* 2014). Other strains of *R. bellii* have been documented to have a native plasmid (Baldridge *et al.* 2010) that could act as a vehicle for the spread of the *tra* genes, and it is conceivable that *R. bellii* RML 369-C may have lost its native plasmid through long term passage in cell culture (Baldridge *et al.* 2008). Therefore, horizontally acquired functional genes and operons in rickettsiae may have been acquired in more

than one way, and while having a native plasmid possibly increases the frequency at which genetic material is spread, it is likely not the only mechanism used. Thus far, the *tra* genes have not been demonstrated to function in bacterial conjugation and may be remnants of a once functional system; however, up-regulation of the *tra* genes in this research implies a potentially, functional conjugative and gene transfer system.

Conclusion

We have demonstrated that the best reference gene out of the ten selected were *metG* or *metG* and *nrdF*. The same reference gene(s) can be used to standardize future relative transcriptional analysis of rickettsiae with similar treatments or to form a basis of potential reference genes with different treatments. Using the identified reference gene(s), we show that the transcription of the *tra* genes of *R. bellii* is highly upregulated at 72 HPI relative to 12 HPI only in tick cell culture, indicating that these genes are very active and contribute to the biology of *R. bellii*. Future research should explore the functions of the *tra* genes guided by bioinformatics to obtain detailed documentation of the mechanism of DNA mobility to and from the chromosome and plasmid, and the demonstration of the process of rickettsial conjugation.

Publication

Chapter 3 was published prior to my dissertation and can be found in this reference (Heu *et al.* 2015).

Fig. 3.1. Graphical display of the domains of *R. bellii* $TraA_{Ti}$ and its related protein. MobA/L: this domain is responsible for specific DNA strand transfer in the donor bacterium; ABC ATPase: functions in ATP and GTP binding and hydrolysis, which is a characteristic of conjugative proteins; UvrD C: this domain is found at the carboxyl terminal end of helicases suggesting their role in helicase activity; Toprim: a domain found in DnaG-type primases; TrwC: this domain signifies the relaxase in conjugative DNA transfer during cell-to-cell contact; Conjugal TraA/I: a protein segment involved in DNA binding, nicking, and helicase activity; Primase: a protein segment involved in priming of DNA. The function of each domain was assigned by the Conserved Domain Database. Rbe = *R. bellii*, Rbu = *R. buchneri*, Atu = *A. tumefaciens*, Eco = *E. coli*, N = amino terminal end of the protein, and C = carboxyl terminal end of the protein.

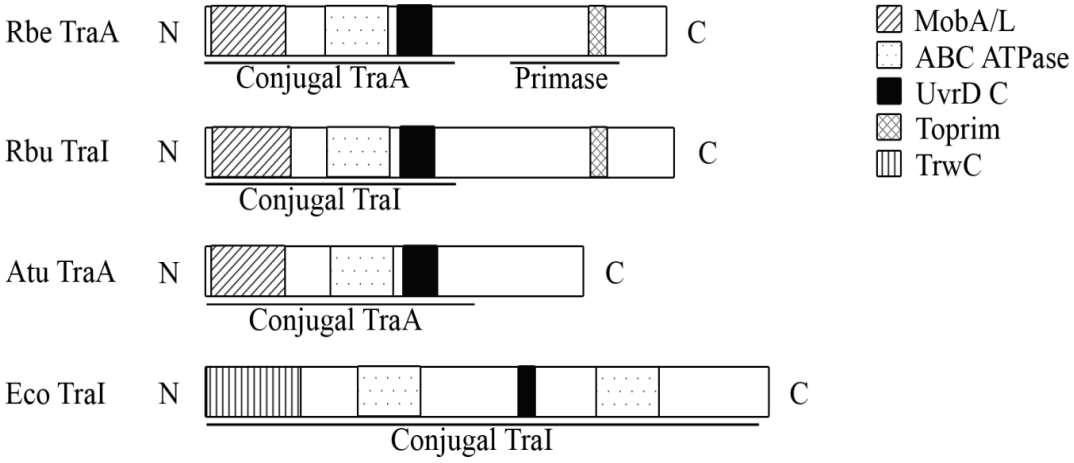


Fig. 3.2. Growth curve of *R. bellii* in ISE6, Vero, and L929. qPCR detected a single copy gene of *gltA* for triplicates of culture at each time point. Standard error bars are shown. Growth trials using Vero and L929 were done simultaneously while ISE6 was done separately.

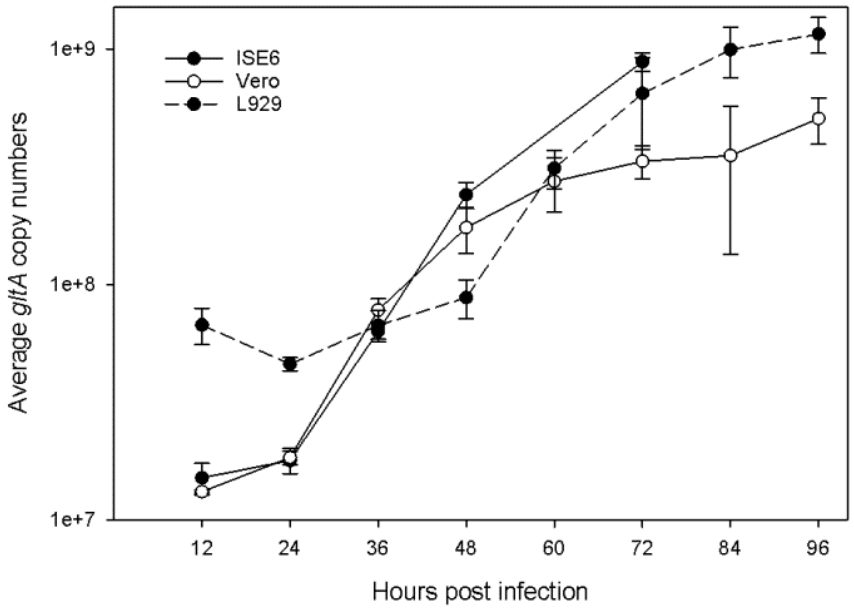


Fig. 3.3. Screenshots of Artemis of *Rickettsia rickettsii* Sheila Smith showing transcription levels of 5 genes: *metG*, *nrdF*, *gyrA*, *gltA*, and 16s rRNA. The order from left to right is of *metG* (A1G_05810), *nrdF* (A1G_03685), *gyrA* (A1G_01555), *gltA* (A1G_07170), and 16s rRNA gene (A1G_r07597). Each vertical black bar represents the hybridization level of mRNA from one probe on the array. The orientation and length of the genes are shown below the transcription peaks.

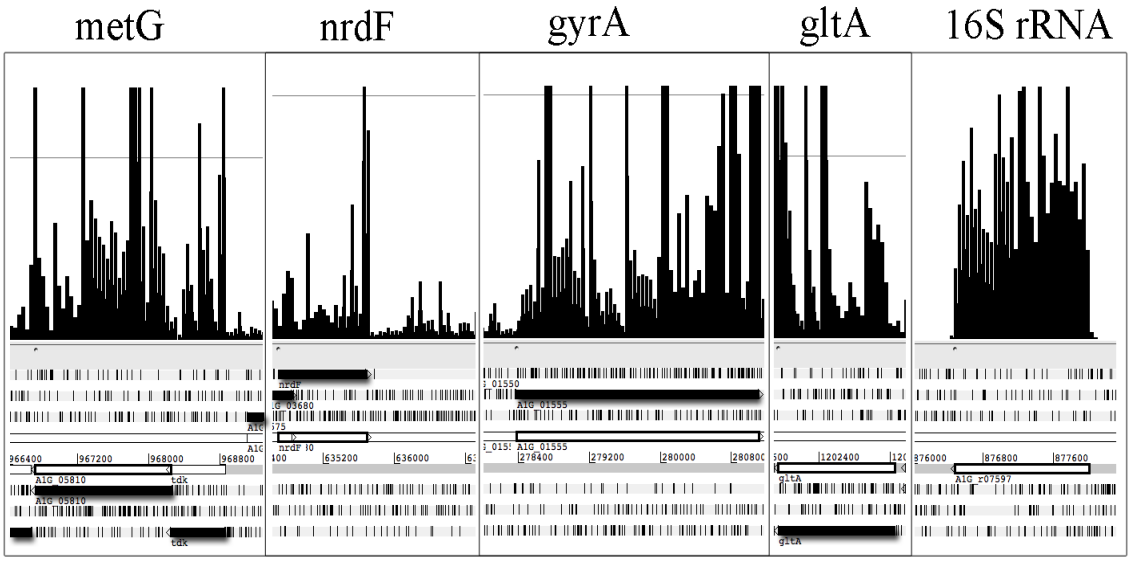


Fig. 3.4. Initial validation of reference genes. *rickA* and *sca2* of *R. bellii* grown in Vero cells were selected to demonstrate differential gene transcription using *metG* as a reference. (**A**) The line graph depicts the fold changes of *rickA* and *sca2* relative to *metG* at all time points to 12 HPI. Letters indicate significant difference within each gene at different times (lower case = *rickA*, upper case = *sca2*) (**B**) The graph shows the transcription level of *rickA* and *sca2* at each time point and (*) indicates significant difference. An ANOVA with Holm-Sidak correction was used to determine significant differences with alpha set at 0.05. Standard error bars are shown.

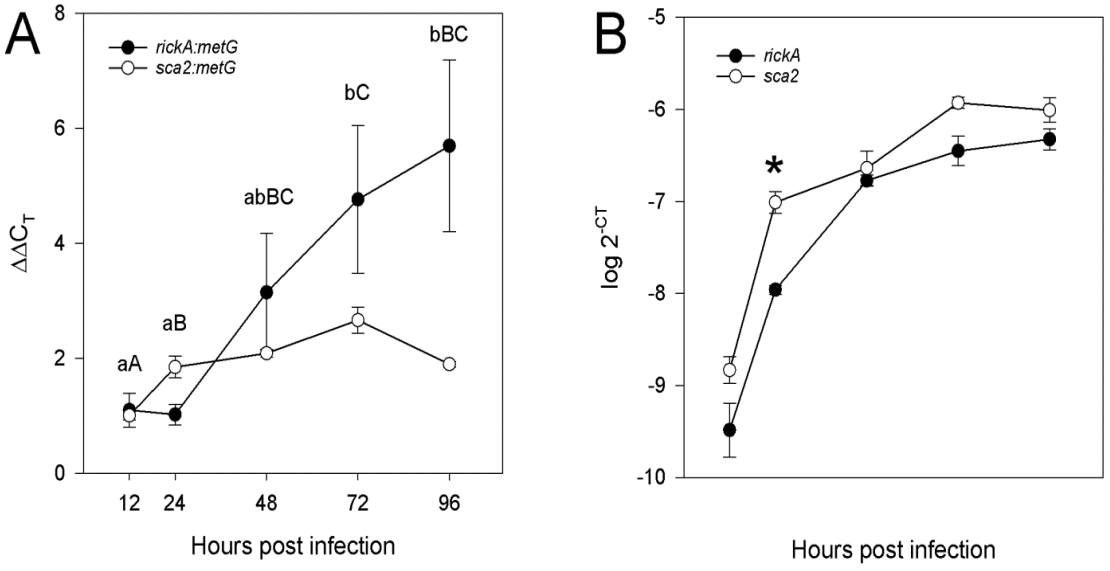


Fig. 3.5. Fold change of *traA$_{Ti}$* in different host cells at different hours post infection. qRT-PCR was used to determine the *traA$_{Ti}$* transcription pattern using *metG* only, *nrdF* only, and a combination of *metG* and *nrdF*. Standard error bars are shown.

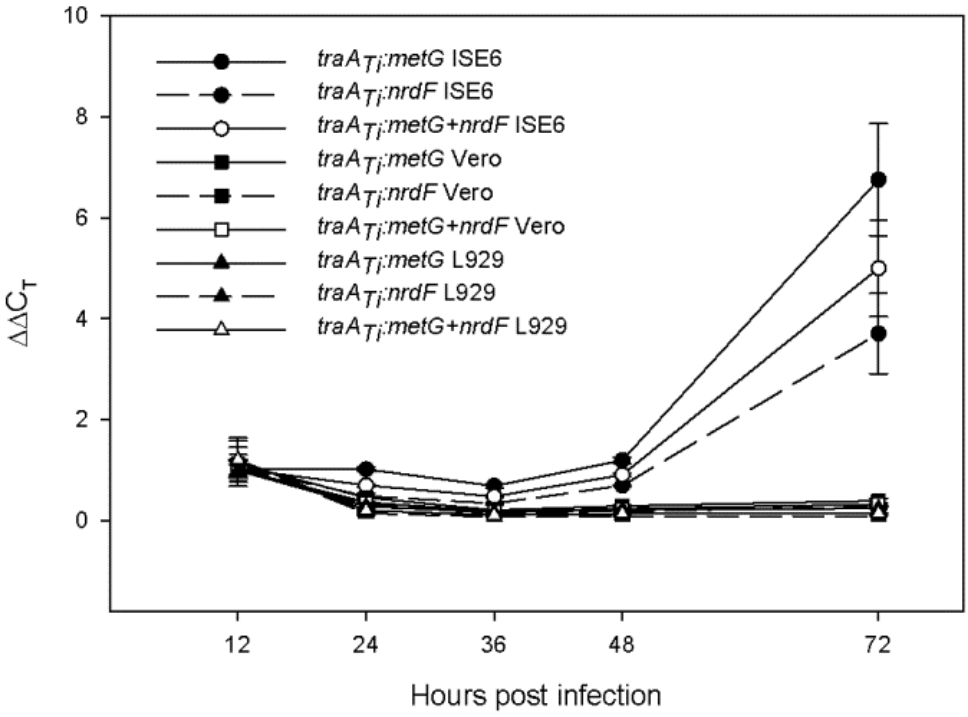


Fig. 3.6. The relative transcription pattern of *metG* and *nrdF* to $traA_{Ti}$. Relative transcription of *metG* and *nrdF* was analyzed using $traA_{Ti}$ as a reference gene to show constitutive, but slightly varying transcription of both genes. Standard error bars are shown.

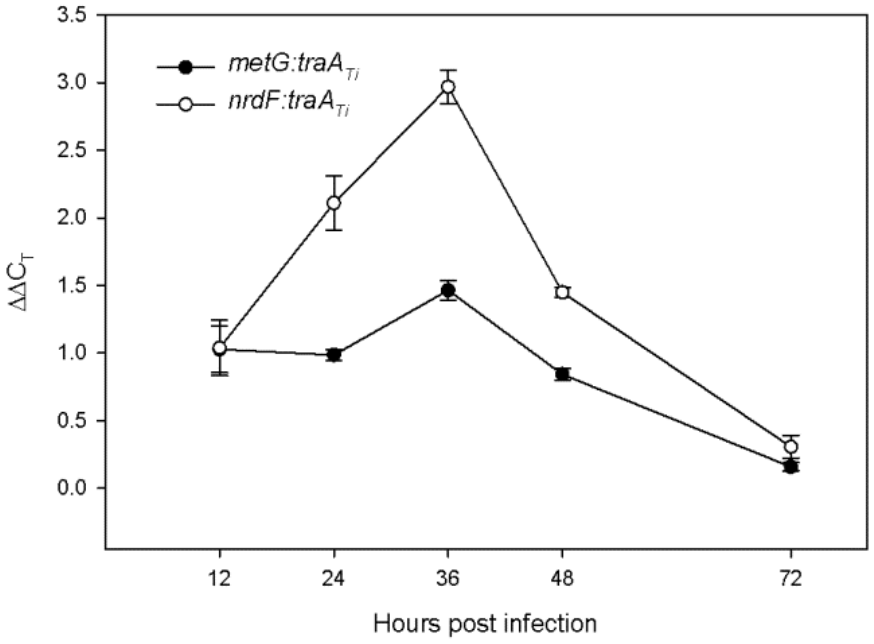


Fig. 3.7. Amino acid alignment of TraA pilin. Sequences include TraA of *Escherichia coli* K-12 (CAA73225.1), TraA of *Enterobacter cloacae* (AKN35276.1), TraA of *Klebsiella pneumoniae* (CDO11547.1), TraA of *Salmonella enterica* subsp. enterica serovar Typhimurium (YP 009077414.1), RBE_0435 of *Rickettsia bellii* RML 369-C (WP 008579911.1), and RMA_0719 of *Rickettsia massiliae* MTU5 (ABV84869.1). Vertical lines indicate cleavage site of the signal peptide predicted by Phobius except for *R. bellii* and *R. massiliae*, which was predicted using Signal-BLAST. Underlined sequences show the transmembrane spanning membrane domain predicted using Phobius. "." = same amino acid above it and "-" = space.

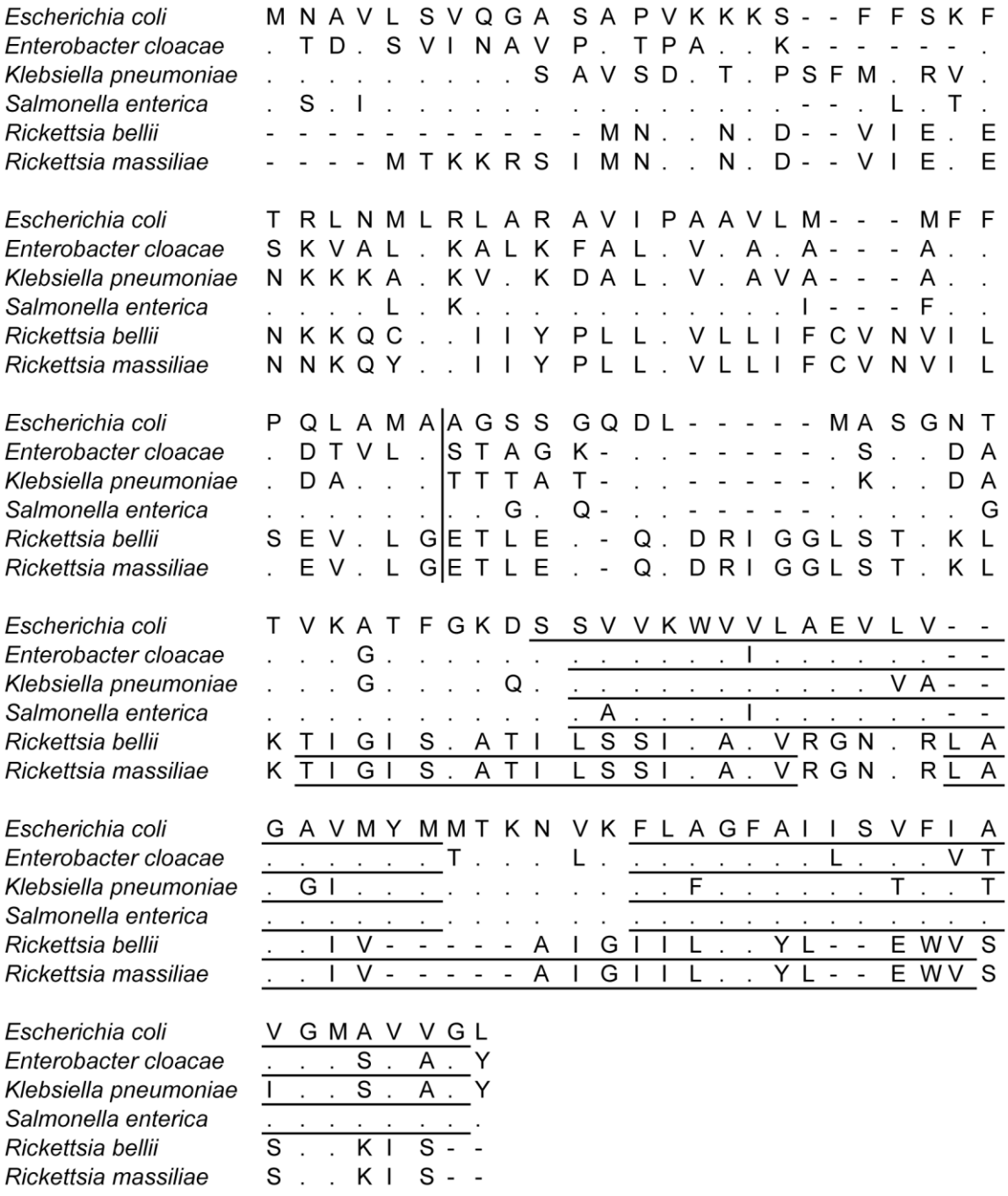

Escherichia coli M N A V L S V Q G A S A P V K K K S - - F F S K F
Enterobacter cloacae . T D . S V I N A V P . T P A . . K - - - - - - .
Klebsiella pneumoniae . . . . . S A V S D . T . P S F M . R V .
Salmonella enterica . S . I . . . . . L . T .
Rickettsia bellii - - - - - - - - - - - M N . . N . D - - V I E . E
Rickettsia massiliae - - - - M T K K R S I M N . . N . D - - V I E . E
Escherichia coli T R L N M L R L A R A V I P A A V L M - - - M F F
Enterobacter cloacae S K V A L . K A L K F A L . V . A . A - - - A . .
Klebsiella pneumoniae N K K K A . K V . K D A L . V . A V A - - - A . .
Salmonella enterica . . . . L . K . . . . I - - - F . .
Rickettsia bellii N K K Q C . . I I Y P L L . V L L I F C V N V I L
Rickettsia massiliae N N K Q Y . . I I Y P L L . V L L I F C V N V I L
Escherichia coli P Q L A M A A G S S G Q D L - - - - - M A S G N T
Enterobacter cloacae . D T V L . S T A G K - . . - - - - - - . S . . D A
Klebsiella pneumoniae . D A . . . T T T A T - . . - - - - - - - . K . . D A
Salmonella enterica . . . . G . Q - . . - - - - - - - . . . . G
Rickettsia bellii S E V . L G E T L E . - Q . D R I G G L S T . K L
Rickettsia massiliae . E V . L G E T L E . - Q . D R I G G L S T . K L
Escherichia coli T V K A T F G K D S S V V K W V V L A E V L V - -
Enterobacter cloacae . . . G . . . . I . . . . . - -
Klebsiella pneumoniae . . . G . . . . Q . . . . V A - -
Salmonella enterica . . . . A . . . . I . . . . . - -
Rickettsia bellii K T I G I S . A T I L S S I . A . V R G N . R L A
Rickettsia massiliae K T I G I S . A T I L S S I . A . V R G N . R L A
Escherichia coli G A V M Y M M T K N V K F L A G F A I I S V F I A
Enterobacter cloacae . . . . . T . . . L . . . . . L . . . V T
Klebsiella pneumoniae . G I . . . . . F . . . . . T . . T
Salmonella enterica . . . .
Rickettsia bellii . . I V - - - - - A I G I I L . . Y L - - E W V S
Rickettsia massiliae . . I V - - - - - A I G I I L . . Y L - - E W V S
Escherichia coli V G M A V V G L
Enterobacter cloacae . . . S . A . Y
Klebsiella pneumoniae I . . S . A . Y
Salmonella enterica . . . . . .
Rickettsia bellii S . . K I S - -
Rickettsia massiliae S . . K I S - -

Table 3.1. Information about the reference/*tra* genes investigated and the qRT-PCR primer pairs

Symbol	Gene Name	Accession #	Function	Primers	Sequence	PCR size
atpB	ATP synthase F0F1 subunit A	ABE05371.1	translocate protons across the membrane	Rbe369C atpB F2	Agttggctttgtaaagcacggcat	154
				Rbe369C atpB R2	Ccatcatatttgccgcaagcctta	
dnaK	Chaperone protein DnaK, Heat shock protein 70	ABE05194.1	folds and unfolds macromolecules	Rbe369C dnaK F2	Tgcatgggttgaagttgaagggaa	163
				Rbe369C dnaK R2	Ttgttgcttgacgctgtgcatcat	
*gltA*1	Citrate Synthase	ABE05406.1	anabolizes citrate for the Kreb cycle	Rbellii qCSF[1]	Ccgcagatgttcacagtgcttt	73
				Rbellii qCSR[1]	Tcatgcatctctttccattgtgc	
gyrA	DNA gyrase subunit A	ABE04875.1	negatively supercoils closed circular DNA	Rbe369C gyrA F1	Gcagcatcaccatccaatgaacca	188
				Rbe369C gyrA R1	Agcaggtaaccatgcaaacagagc	
infB	Translation initiation factor IF-2	ABE04694.1	initiates protein translation	Rbe369C infB F1	Acgctacctgaagttattggtg	189
				Rbe369C infB R1	Caacatccgattcttggactc	
metG	Methionyl-tRNA synthetase	ABE04319.1	ligates methionine to tRNA	Rbe369C metG F1	Tatggcacatggctggtggactaa	166
				Rbe369C metG R1	Gattgctgcgggcaaagttaccat	
nrdB	Ribonucleotide-diphosphate	ABE04813.1	catalyzes ribonucleotides to	Rbe369C nrdF F1	Tttcatcacgcactgaccaagcaa	136
				Rbe369C nrdF R1	Aactacgcttgcagtatttggagc	

	reductase subunit beta		deoxyribonucleotides			
rpoB	DNA-directed RNA polymerase subunit beta	ABE05234.1	catalyzes the transcription of DNA to RNA	Rbe369C rpoB F2	Tgatgcccaaccaagatgag	254
				Rbe369C rpoB R2	Caaagcggagacgatttacc	
tlc5	ADP/ATP translocase 5	ABE04222.1	exchanges ADP from rickettsia with ATP from the host	Rbe369C tlc5 F2	Attggatggaacgctacccgcatt	99
				Rbe369C tlc5 R2	Atattaggccaaagctcggcaagg	
16s rRNA	16s ribosomal RNA	ABE04554.1	smaller subunit of a ribosome responsible for protein synthesis	qPCR16SF2	Aaggcggtcatctgggctacaac	97
				qPCR16SR2	Cactcatcgtttacggcgtggac	
rickA	Actin polymerization protein RickA	ABE04936.1	recruits Arp2/3 complex for actin tail motility	qRbRickA F2[2]	Tacgccactccctgtgtca	146
				qRbRickA R[2]	Gatgtaacggtattatcaccaacag	
sca2	Cell surface antigen Sca2	ABE05361.1	polymerizes unbranched actin tail	qRbe sca2 F	Aggacgcaataaacagaggag	115
				qRbe sca2 R	Cataccgatagttacaccgct	
traATi	Conjugal transfer	ABE04502.1	3	traAti qpcr F	Agtagatgaagcagggatggtaggt	131

	protein TraATi			traAti qpcr R	Aacattcctcctcgctcaacgg	
traDF	Conjugative transfer protein TraDF	ABE04504.1	3	traDf F	Attggttgccttggaatgattgtc	120
				traDf R	Gcgagtaccaactcagcactt	
Rbe_0422	Transposase	ABE04503.1	3	qRbe0422 F	Cattgacgaagatagagcaagc	175
				qRbe0422 R	Ttggatacctttagcctcacc	
traL	Conjugative transfer protein TraL	ABE04519.1	3	RbeqPCR TraL F	Ttggtcactctctactcc	102
				RbeqPCR TraL R	Ccgctattcagtaacactat	
traE	Conjugative transfer protein TraE	ABE04518.1	3	RbeqPCR TraE F	Cagttacgagtcttcttgg	76
				RbeqPCR TraE R	Ggctctattgctggaatta	
traB	Conjugative transfer protein TraB	ABE04515.1	3	RbeqPCR TraB F	Cccgaggtttacaagaag	76
				RbeqPCR TraB R	Catttagcccttggatctat	
traV	Conjugative transfer protein TraV	ABE04514.1	3	RbeqPCR TraV F	Gatcagaatatgtttaatcca	83
				RbeqPCR TraV R	Ataccttccaaagcatct	
traC	Conjugative transfer protein TraC	ABE04513.1	3	RbeqPCR TraC F	Gatttagcaggagcagtag	99
				RbeqPCR TraC R	Ccctttggctttgttct	
traW	Conjugative transfer protein TraW	ABE04512.1	3	RbeqPCR TraW F	Aagcaggaaacatcatagta	99
				RbeqPCR TraW R	Gctcatcatcaccatcaa	

traU	Conjugative transfer protein TraU	ABE04511.1	3	RbeqPCR TraU F	Cattcctatgccaattcct	98
				RbeqPCR TraU R	Caccaaggcttaccatac	
trbC	Conjugative transfer protein TrbC	ABE04510.1	3	RbeqPCR TrbC F	Ggcacaaccagaagtatta	106
				RbeqPCR TrbC R	Tgctgaatcaggcattg	
traN	Conjugative transfer protein TraN	ABE04509.1	3	RbeqPCR TraN F	Caatggcacaaagattaca	94
				RbeqPCR TraN R	Gacgattccctcctaaac	
traF	Conjugative transfer protein TraF	ABE04508.1	3	RbeqPCR TraF F	Cgcatttgctggtactat	101
				RbeqPCR TraF R	Ttggttgagtgctatcttt	
traH	Conjugative transfer protein TraH	ABE04507.1	3	RbeqPCR TraH F	Cagcaggagggttatct	88
				RbeqPCR TraH R	Ccactacagctcatgttta	
traG	Conjugative transfer protein TraG	ABE04506.1	3, 4	RbeqPCR TraG F	Ctggtaaaggcacaagtc	112
				RbeqPCR TraG R	Ggaactgaagctgctaac	
traDTi	Conjugative transfer protein TraDTi	ABE04501.1	3	RbeqPCRTraDti F	Ctcgtactcgtcgtttaat	78
				RbeqPCRTraDti R	Gtaaactgttagtgggtaaatc	

1 = (Baldridge *et al.* 2010), 2 = (Oliver *et al.* 2013), 3= (Ogata *et al.* 2006), 4= (Audette *et al.* 2007), nd= not determined

Table 3.2. "Area under the curve" from RrSS tiling microarray for five genes selected

Gene	**Locus**	**Mean of triplicate counts of "area under the curve"**
16S rDNA	A1G_r07597	10858949
gltA	A1G_07170	17600
gyrA	A1G_01555	95527
metG	A1G_05810	23697
nrdF	A1G_03685	4827

Table 3.3. A comparison of *gltA* copy numbers of RP and WC *R. bellii*

Time	12 HPI		72 HPI	
Method	RP	WC	RP	WC
Mean	$4.01x10^7$	$3.77x10^7$	$3.72x10^9$	$4.13x10^9$
S.E.	$6.98x10^6$	$1.05x10^7$	$2.04x10^8$	$2.62x10^8$
$p \leq 0.05$	0.86		0.28	

Table 3.4. The rankings of each candidate gene by Normfinder and BestKeeper

	Normfinder		**BestKeeper**	
	Stability Value		**Coefficient of Correlation**	
Gene name	**Different HPI in ISE6**	**Different host cells**	**Different HPI in ISE6**	**Different host cells**
16S rRNA	0.233	0.471	0.916	0.872
atpB	0.216	0.314	0.923	0.918
dnaK	0.129	0.116	0.983	0.973
gltA	0.314	0.373	0.993	0.978
gyrA	0.157	0.093*	0.995**	0.987**
infB	0.212	0.227	0.965	0.960
metG	0.080**	0.080**	0.994*	0.983*
nrdF	0.117*	0.211	0.990	0.952
rpoB	0.232	0.373	0.920	0.904
tlc5	0.315	0.147	0.921	0.981

**= most stable reference gene, *= second most stable reference gene

Table 3.5. C_T values of *tra* genes and *metG* at 12 and 72 HPI

***tra* genes**	***metG* 12 HPI**	***metG* 72 HPI**	***tra* 12 HPI**	***tra* 72 HPI**	**ddC_T**
traL	27.65±0.05	25.44±0.23	33.49±0.16	28.91±0.25	**5.12±0.08**
traE	27.65±0.05	25.44±0.23	36.76±0.35	31.98±0.38	**6.01±0.64**
traB	25.85±0.06	23.7±0.27	33.64±0.26	28.42±0.18	**8.45±0.53**
traV	27.01±0.06	24.98±0.21	36.38±0.13	31.19±0.26	**8.94±0.46**
traC	25.85±0.06	23.7±0.27	35.63±0.09	30.27±0.21	**9.39±0.68**
traW	25.85±0.06	23.7±0.27	30.76±0.12	25.92±0.21	**6.48±0.27**
traU	27.65±0.05	25.44±0.23	36.92±0.02	31.21±0.34	**11.35±0.83**
trbC	27.65±0.05	25.44±0.23	35.76±0.12	30.01±0.23	**11.20±0.47**
traN	27.19±0.05	24.9±0.25	33.75±0.24	28.9±0.29	**5.92±0.17**
traF	27.19±0.05	24.9±0.25	ND	35.49±0.20	**NA**
traH	27.19±0.05	24.9±0.25	34.2±0.11	29.39±0.22	**5.80±0.35**
traG	27.19±0.05	24.9±0.25	32.91±0.05	26.37±0.20	**19.20±0.91**
$traD_F$	26.56±0.02	24.2±0.31	26.69±0.03	23.26±0.25	**2.11±0.08**
$traA_{Ti}$	26.56±0.02	24.2±0.31	31.81±0.13	27.55±0.23	**3.75±0.21**
$traD_{Ti}$	25.32±0.08	23.11±0.15	36.53±0.15	31.56±0.22	**6.81±0.38**
Rbe_0422	30.47±0.14	24.36±.27	31.63±0.08	25.20±0.36	**1.25±0.08**

C_T values are given with standard errors for *metG* and *tra* genes at 12 and 72 HPI. ND= No detection, NA= Not available.

Chapter 4: Complementation assay using *traD*$_F$ of *Rickettsia bellii* in a *traD* mutant *Escherichia coli* genetic background

Introduction

The genomes of *Rickettsia*, obligate intracellular bacteria, are being sequenced at a fast rate because of current technological advances. Genome comparison and bioinformatics have shown that conjugation genes, called the transfer or *tra* genes, are common in *Rickettsia* (Weinert, Welch and Jiggins 2009), suggesting that they may be functional genes and that *Rickettsia* may be proficient in horizontal transfer of genetic information. The origin of the rickettsial *tra* genes has been hypothesized to derive from *Legionella* spp. (Ogata *et al.* 2006) during coinfection of an amoeba. These rickettsial genes are related to the *E. coli* F-plasmid *tra* genes and the *A. tumefaciens* tumor-inducing *virB/D* genes. Bioinformatic analysis supports the conclusion that these rickettsial *tra* genes belong to the bacterial conjugation and mobilization family.

However, functional analyses of the *tra* genes have not been done. Assays complementing bacterial *tra* mutants using rickettsial homologs to restore function would support the conclusion that these genes function as bacterial conjugation genes. The *traD*$_F$ gene that encodes the predicted coupling protein of the F-type IV secretion system is the most likely candidate of the *tra* genes of *R. bellii* that may complement a *traD* mutant *E. coli* based on other complementation assays of coupling proteins from different conjugation systems (Hamilton *et al.* 2000). Coupling proteins have two transmembrane-spanning domains (at the N-terminal end in most coupling proteins) and a water soluble portion (at the C-terminal end in most proteins) that can sometimes determine specificity

of the relaxosome (Panicker and Minkley Jr 1992; Sastre, Cabezon and de la Cruz 1998). The interaction of the relaxosome and the coupling protein is an essential step in bacterial conjugation although the mechanism and role of the coupling protein is still in question (Gomis-Ruth *et al.* 2004). Mutation in the coupling protein nullifies or greatly decreases the efficiency of DNA transfer (Walker and Pittard 1969; Moncalian *et al.* 1999). Interestingly, some coupling proteins can recognize more than one relaxosome and can interact with the mating pair formation of other origins, e.g. TraG of RP4 allows mobilization of plasmid RSF1010, a nonself-transmissible, broad range host plasmid, in *A. tumefaciens* (Hamilton *et al.* 2000). In fact, the first complementation of two different classes of coupling proteins was determined by substituting *trwB* of R388 with *traG* of RP4 to mobilize RSF1010 in *E. coli* (Cabezon, Lanka and de la Cruz 1994). Interestingly, the last 37 amino acids of TraD of *E. coli* were shown to be specific for TraI of *E. coli*; however, the deletion of this segment allowed recognition of the R388 relaxosome (Sastre, Cabezon and de la Cruz 1998). *Rickettsia bellii* $TraD_F$ is about 150 amino acids shorter than TraD of *E. coli* and may already be missing the selective carboxyl-end. Coupling proteins are one of the conjugation constituents that are most conserved (Gomis-Ruth *et al.* 2004). Using *E. coli* as a model bacterium for bacterial conjugation, it is possible that hypothetical conjugation genes can be characterized especially for a more conserved, essential protein like the coupling protein.

A complementation assay was designed using the bacterial conjugation machinery of *E. coli* as a phenotype to characterize $traD_F$ of *R. bellii* in an *E. coli traD* mutant genetic background. The experimental design is comprised of two *E. coli* donors and

recipients. The donors consist of a strain of *E. coli* with its normal codons (BL21(DE3)) and a codon-supplemented strain of *E. coli* (Rosetta 2(DE3) pLysS). The plan was to transform both strains with a derivative of the F-plasmid (F+SG or F-*traD36*+SG plasmid) and pet28A-$traD_F$ or -*traD* gene from *R. bellii* (unknown) or *E. coli* (positive control), respectively. Donor strains were transformed with pet28A-$traD_F$ or -*traD* to screen for the expression of $TraD_F$ or TraD. Subsequently, these transformants could be transformed again with a construct consisting of the *tra* genes from the wild-type F-plasmid or F-*traD36* plasmid from *E. coli* K-12 or JM109, respectively, with an oriT (origin of transfer), oriV (origin of replication) and a selectable marker for this plasmid (spectinomycin resistance). These plasmids were expected to provide the necessary genetic background to screen for the complementation of *traD36* by $traD_F$, as well as wild-type *traD* for a positive control of bacterial conjugation. The other component of this experimental design was to choose recipients that could be selected for. The recipients consisted of *E. coli* strains BZB1203 and HMS180 with internal antibiotic resistance markers (nalidixic acid and rifampicin resistance, respectively) that could be selected against the donors and for the recipients. The recipients could be double selected for the native antibiotic resistance and for spectinomycin to select for the transfer of F-plasmid or F-*traD36* plasmid. Therefore, successful isolation of the double-selected recipient would demonstrate the complementation of *traD36* by $traD_F$.

In this experiment, the F+SG and F-*traD36*+SG plasmids were successfully constructed. BL21(DE3) and Rosetta 2(DE3) pLysS were transformed with pet28A-$traD_F$ or –*traD*; however, $TraD_F$ was not expressed in BL21(DE3) or Rosetta 2(DE3)

pLysS despite the various conditions tested. TraD was shown to express in the *E. coli* strains transformed with pet28A-*traD* in all the conditions tested. This controlled for the *E. coli* strains and expression plasmid vector used in this experiment. The lack of $TraD_F$ expression prevented the completion of the complementation assay and future directions are discussed in chapter 5 of this thesis.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains used were *R. bellii* RML 369-C, *E. coli* JM109 (e14– (McrA–) *recA1 endA1 gyrA96 thi-1 hsdR17 (rK – mK +) supE44 relA1* Δ*(lac-proAB)* [F′ *traD36 proAB lacIq ZΔM15*]) (Promega, WI), *E. coli* K-12 wild type (*ΔrfbB51*) (Liu and Reeves 1994), *E. coli* BZB2103 (F-, λ-, *gyrA586*(NalR), IN(*rrnD-rrnE*)1, *rph-1* [pColD-CA23]) (Pugsley 1985), and *E. coli* HMS180 (F-, λ-, *thyA36*, IN(*rrnD-rrnE*)1, *rph-1*?, *rpoB331*(rifR), *hsdR19*) (Campbell, Richardson and Studier 1978). *Rickettsia bellii* was cultured in *Ixodes scapularis* embryonic cells, ISE6, in 25 cm^2 tissue culture flasks (BD Falcon, NJ) at 34°C in L15-C300 supplemented with 10% FBS, 5% tryptose phosphate broth, and 0.1% lipoprotein concentrate (Munderloh and Kurtti 1989; Oliver *et al.* 2013). JM109, BZB2103, HMS180, and K-12 were cultured in 2x yeast extract/tryptone (YT) medium at 37°C, with shaking at 250 revolutions per minute (rpm), over night with nalidixic acid [50 µg/µL] for JM109 and BZB2103, rifampicin [100 µg/µL] for HMS180, or no antibiotics for K-12. Glycerol stocks were made for each *E. coli* strain by adding a 1:1 ratio of the overnight culture to 50% glycerol in a cryotube

and storing at -70°C. All *E. coli* strains except JM109 were provided by Coli Genetic Stock Center (CGSC) at Yale University.

Cloning of *traD*$_F$ of *R. bellii* and *traD* of *E. coli* K-12 into pET28A

Rickettsia bellii were cultured in ISE6 cells in L15C-300 supplemented with 10% fetal bovine serum, 5% tryptose phosphate broth, 0.1% lipoprotein concentrate to generate DNA template for the cloning of *traD*$_F$. *Rickettsia bellii*-infected ISE6 cells were scraped off the flask, suspended in L15C-300 complete medium, transferred to 1.5 ml microcentrifuge tubes containing 60-90 grit silicone carbide (Lortone, Inc., Mukilteo, WA), and vortexed at maximum speed for at least 20 seconds to rupture host cells and obtain cell-free *R. bellii*. Liberated *R. bellii* were filter-purified using a Whatman 2.0 µm filter (GE Healthcare Life Science, NJ) as described in Oliver *et al.* (2014), and collected by centrifugation at 13,000x rcf for 5 minutes at 4°C. *Escherichia coli* K-12 was cultured as previously mentioned and centrifuged at 13,000x rcf for 5 minutes at 4°C. DNA Blood and Tissue kit (QIAGEN, MN) was used to extract the DNA of *R. bellii* and *E. coli* K-12 following the manufacturer's protocol for extracting the DNA of Gram-negative bacteria. Subsequently, *traD*$_F$ (*R. bellii*) and *traD* (*E. coli*) were cloned from PCR products using Q5 high-fidelity hot start polymerase (New England BioLab, MA), primers listed in Table 4.1 at [1 µM], deoxynucleotide mix [0.4 mM], and DNA template [2 µg/µL]. Thermal cycles for amplifying *traD*$_F$ or *traD* were 1 cycle at 94°C for 30 seconds; 2 cycles at 94°C for 10 seconds, 55 or 62°C, respectively, for 1 minute, and 72°C for 1 minute; 35 cycles at 94°C for 10 seconds, 67 or 72°C, respectively, for 1 minute, and 72°C for 1 minute; and 1 cycle at 72°C for 2 minutes.

The *traD* PCR products were cloned as follows: PCR reactions were purified using the Clean and Concentrator kit (Zymo Research, CA) and restriction enzyme digested with EcoRI-SacI or EcoRI-SalI for $traD_F$ or *traD*, respectively, along with pet28A (Novagen, MA), a protein expression vector, according to the recommended NEB buffer and BSA at 1x for 1.5 hours at 37°C. Subsequently, restriction enzymes were inactivated by incubating reactions at 65^{o}C for at least 15 minutes. pet28A was incubated with Antarctic phosphatase (New England BioLab, MA) according to the manufacturer's protocol to dephosphorylate the 5' end to prevent self-ligation. All reactions were cleaned using the Clean and Concentrator kit and ligated using T4 ligase (New England BioLab, MA) at 15°C overnight with a 1:3 ratio of pET28A: $traD_F$ or *traD*. Ligations were used to transform ElectroMAX Stbl4 competent cells (Invitrogen, NY). Cells were plated on 2x YT agar with kanamycin [50 µg/mL] and incubated at 37°C overnight. Colonies were cultured overnight and plasmids harvested using the High Pure Miniprep kit (Roche, IN). Plasmids were screened for inserts by restriction enzyme digestion with a combination of enzymes and conditions as mentioned above, electrophoresed in a 1% agarose gel for 1.5 hours at 80 V, and stained with 1x Gel Green dye (Biotium, CA) for 1 hour. Bands were visualized by using a dark reader (Clare Chemical Research, CO) and an image was taken in ImageJ v. 1.23. Construct sequences were verified using Sanger sequencing to ensure that the $traD_F$ and *traD* reading frames were aligned with the pet28A reading frames and to verify that the sequences had no mutations.

Diagrams of the plasmids were constructed in MacVector (MacVector, Inc., NC) using sequences of pET28A and $traD_F$ (NC_007940.1, region: 476440-478128) of *R. bellii* or *traD* (NC_002483, region: 89804-91957) of *E. coli* K-12.

Constructing F+SG and F-*traD36*+SG

Escherichia coli K-12 and JM109 were cultured as previously mentioned and centrifuged at 13,000x rcf for 5 minutes. The plasmids were extracted from the bacterial pellets through alkaline lysis using the resuspension buffer, lysis buffer, and protein precipitation buffer provided in the High Pure Plasmid kit. Pellets of the *E. coli* strains were resuspended, lysed, and the proteins were precipitated using the buffer provided by High Pure Plasmid kit. After pelleting the proteins at 13,000x rcf for 10 minutes, F- or F'-*traD36* plasmids in solution were transferred to a heavy phase lock gel 2.0 mL tube (5 PRIME, MD). An equal volume of phenol:chloroform:isoamyl (Sigma-Aldrich, MO) was added to the plasmid solution, mixed, and centrifuged at 13,600x rcf for 3 minutes. The aqueous phase was transferred to a new heavy phase lock gel and subjected to the same treatment and an additional treatment of only chloroform. Aqueous phase was transferred to a new centrifuge tube and sodium acetate was added to a final concentration of [0.3 M] to help with precipitating the DNA. Two volumes of ice-cold 100% ethanol was added, mixed, incubated on ice up to 1 hour, and centrifuged at 13,600x rcf for 3 minutes to recover the plasmids. F- and F-*traD36* plasmids were digested with HindIII (New England BioLabs, MA) following the manufacturer's protocol, dephosphorylated to prevent self-ligation, and cleaned using the Genomic DNA Clean and Concentrator kit (Zymo Research, CA).

The spectinomycin resistance marker (Spec) and green fluorescent reporter gene (GFPuv), hereafter referred to as the SG fragment, were from pRAM18dSGK#23[MCS] #50 (Burkhardt *et al.* 2011) and referred to as plasmid B in this chapter. Plasmid B was digested using HindIII, electrophoresed in a 0.8% agarose gel at 80 V for 1.5 hours, stained in 1x Gel Green and visualized using a dark reader. The band corresponding to the predicted size of the SG fragment was extracted using the DNA gel extraction kit (Zymo Research, CA).

The SG fragment and the HindIII-digested F- or F-*traD36* plasmid were ligated at a ratio of 8:1 ratio, respectively, using T4 ligase at 15°C overnight. The ligation was used to transform Stbl4 competent cells and *E. coli* maintaining the plasmids were selected using spectinomycin [100 µg/mL]. Plasmids containing the minimal fragment necessary for bacterial conjugation and plasmid replication of F-plasmid or F-*traD36* with SG fragments were termed F+SG or F-*traD36*+SG, respectively. F+SG and F-*traD36*+SG were subjected to HindIII digestion and pulsed-field gel electrophoresed using the CHEF Mapper XA system (BIO-RAD, CA) with conditions set at 1% agarose gel, 120° angle, 60 V, linear ramping, and an initial and final switch time of 0.14 and 8.53 seconds; respectively. The 5 kb (BIO-RAD, CA) and 1 kb+ (Invitrogen, NY) linear DNA ladder were used to determine the sizes of the bands.

F+SG and F-*traD36*+SG were verified to contain the complete *tra* cluster of the F-plasmid using PCR to detect *traM* (the first gene of the *tra* cluster) and *traD* (the third to the last gene of the *tra* cluster and the gene of interest).

Expression of $TraD_F$ and TraD in BL21(DE3) and Rosetta 2(DE3) pLysS

BL21(DE3) (New England BioLabs, MA) and Rosetta 2(DE3) pLysS (Novagen, MA) were transformed with pet28A-*traD*$_F$ or –*traD* that were sequenced to verify integrity and alignment of the reading frames. Three single colonies were screened for the expression of TraD$_F$ and TraD. In brief, colonies were cultured in 2x YT media in a 14 mL round-bottom tube (BD Falcon, NJ) overnight at 37°C with shaking at 250 rpm. Five mL of fresh 2x YT medium in a new 14 mL round-bottom tube were inoculated with the overnight culture at a 1:50 ratio (v/v) and cultured at 37°C until an OD600 of 0.6-8 was reached. Cells were centrifuged at 1,500x rcf for 5 minutes, supernatant was discarded, and cell pellets were resuspended in 5 mL of M9 minimal medium with glycerol, superior broth (AthenaES, MD), or in 2x YT medium. Cultures were grown at 37^{o}C and 250 rpm for an additional 30 minutes to adapt the bacteria to the new medium. Bacterial cultures were induced with 0, 0.1, or 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 15, 30, or 37°C with shaking at 200 rpm. Cultures were sampled overnight or at 4 hours post induction for cultures induced at 15 or 30 and 37°C, respectively.

To check for the expression of TraD$_F$ and TraD, samples of induced or uninduced *E. coli* culture were aliquoted into microcentrifuge tubes. Bacteria samples were centrifuged at 13,600x rcf for 5 minutes and the supernatant was discarded. 6x gel loading buffer (11 mM EDTA, 3.3 mM Tris-HCl, 0.017% SDS, 0.015% bromophenol blue, pH 8.0 at 25°C) was used to resuspend the bacterial pellets, boiled for 10 minutes to lyse bacteria and denature the protein, loaded onto a 10-well 10% Mini-PROTEAN® TGX™ Precast Protein Gel (BIO-RAD, CA) in 1x Tris-glycine buffer, and electrophoresed at 90 V for 1 hour. Polyacrylamide gels were subsequently stained in a

modified Coomassie Brilliant Blue staining solution (Lawrence and Besir 2009) or probed for His-tagged protein in a Western blot. In brief, the polyacrylamide gel was boiled in two changes of water to remove SDS. The water was discarded and the gel was boiled briefly in Coomassie Brilliant Blue G-250 solution (Lawrence and Besir 2009) for 5-10 minutes, rinsed and stored in water. For Western blotting, proteins were transferred to a nitrocellulose membrane using a Mini Trans-Blot® Cell (BIO-RAD, CA) with 20% methanol in 1x Tris-glycine buffer at 30 V overnight at 4°C. The membrane was blocked in 5% non-fat dry milk, rinsed with 1x phosphate buffered saline (PBS) three times, and probed with anti-His (HIS.H8/ ThermoFisher Scientific, NY) monoclonal antibody derived from mouse (mab) (1:2000). After three rinses with PBS, membranes were incubated with horseradish peroxidase-conjugated anti-mouse goat antibody (1:5000), washed with PBS, and protein bands were detected using the Pierce™ DAB substrate kit (ThermoFisher Scientific, NY). In addition, the pet28A-*traD*$_F$ or –*traD* were reisolated from BL21(DE3) and Rosetta 2(DE3) pLysS and verified to have no mutations or misalignment in the reading frames.

Results

Plasmid construction and isolation

The pet28A-*traD*$_F$ and –*traD* were successfully constructed using the Q5 polymerase chain reaction to amplify *traD*$_F$ and *traD* having EcoRI-SacI or EcoRI-SalI restriction enzyme sites, respectively, and ligated with pet28A digested with the same restriction enzymes. Stbl4 cells were transformed with this ligation and colonies were picked and screened for correct DNA sizes following restriction enzyme digestion. This

construct was confirmed to have the correct predicted sizes for pet28A (~5,300 bp), *traD*$_F$ (~1,600 bp), or *traD* (~2,100 bp) (Fig. 4.1B). Sequencing confirmed the integrity of *traD*$_F$, *traD*, and the T7 promoter of pet28A. A general graphical representation of both pet28A-*traD*$_F$ and –*traD* is depicted in figure 4.1A.

The F+SG and F-*traD36*+SG plasmids were constructed using HindIII-digested fragments of the F-plasmid from *E. coli* K-12 WT and F'-*traD36* plasmid from JM109 (Fig. 4.2A), respectively, and ligated with the SG-fragment, or spectinomycin and green fluorescent protein fragment (Fig. 4.2B). This ligation was transformed into Stbl4 cells and colonies were screened for the construct with the correct sizes using restriction enzyme digestion of HindIII to detect DNA sizes of ~55 kb for F or F'-*traD36* backbone and ~3.1 kb for the SG-fragment (Fig. 4.2C). The presence of the *tra* cluster in F+SG and F-*traD36*+SG were confirmed through PCR detection of *traM* and *traD* (Fig. 4.2D).

Protein expression of pet28A-*traD*$_F$ and –*traD*

BL21(DE3) and Rosetta 2(DE3) pLysS were transformed with pet28A-*traD*$_F$ and –*traD* and induced with 0, 0.1, or 0.4 mM IPTG for expression at 15, 30, or 37°C with shaking at 200 rpm in M9 minimal medium, superior broth, and 2x YT medium. Coomassie Brilliant Blue staining of polyacrylamide gels revealed recombinant protein bands in BL21(DE3) and Rosetta 2(DE3) pLysS transformants (Fig. 4.3 A/B and 4.4 A/B). In addition, a more intensely stained band migrating near the 95 kDA was apparent in the BL21(DE3) transformants with induction of 0.1 and 0.4 mM IPTG and cultured in all three media and temperature (Fig. 4.3A). Western blot analysis of the total proteins of BL21(DE3) and Rosetta 2(DE3) pLysS transformants identified multiple protein bands in

the pet28A-*traD* transformants after IPTG induction (Fig. 4.3C and 4.4C). One band migrating just above the 95 kDa marker and another between the 72 and 95 kDa may potentially be TraD+His-tag (estimated size of TraD is approximately 85 kDa). In contrast, transformants containing pet28A-$traD_F$ displayed a few new protein bands that did not correlate with the estimated sizes of $TraD_F$+His-tag (~ 68 kDa) in BL21(DE3) transformants (Fig. 4.3D), and no new protein bands were detected in Rosetta 2(DE3) pLysS transformants (Fig. 4.4D).

Discussion

This experiment was designed to test if $traD_F$ of *R. bellii* was able to complement a *traD* mutant of *E. coli* in bacterial conjugation. This consists of constructing plasmids that expressed $TraD_F$ and TraD proteins (pet28A- $traD_F$ and – *traD*) in a potential *traD* mutant *E. coli* donor (BL21(DE3) or Rosetta 2(DE3) pLysS transformants) and allowing these transformants to transfer a selectable plasmid (spectinomycin resistance on the F-SG or F-*traD36*+SG plasmid) to a recipient (BZB1203 or HMS180) that can then be doubly selected.

The plasmid constructs were successfully generated and verified through restriction endonuclease digestion with the corresponding enzymes (Fig. 4.1B, 4.2C), sequencing of pet28A-$traD_F$ and –*traD,* and PCR detection of the *tra* cluster in F+SG and F-*traD36*+SG (Fig. 4.2D). Unfortunately, expression of the pet28A-$traD_F$ in BL21(DE3) cultured and induced at different concentration of IPTG, at different temperatures, and in different media yielded no protein bands that corresponded to $TraD_F$, but only to the control, TraD (Fig. 4.3A-D). The smaller size protein bands detected on Western blot in

BL21(DE3) transformed with pet28A-*traD*$_F$ were detected in the 0, 0.1, and 0.4 mM IPTG induced cultures (Fig. 4.3D), as well as the no induction in BL21(DE3) transformed with pet28A-*traD* cultured in M9 minimal and 2x YT media (Fig.4.3C). Collectively, this suggested that BL21(DE3) expressed native proteins with sequential histidines that corresponded to 34, 26, and 17 kDa that were detected by anti-His-tag antibodies (Fig. 4.3C/D). The Rosetta 2(DE3) pLysS transformed with pet28A-*traD*$_F$ showed no signs of His-tagged proteins, which is surprising. This strain carried a plasmid that supplemented rare tRNAs in *E. coli* and should have been able to accommodate the AT-rich codons of the rickettsial *traD*$_F$ gene. Neither different temperatures nor IPTG concentrations, or media favored the expression of TraD$_F$, but favored the expression of TraD acting as a positive control for the bacterial strains and expression vector.

The AT-rich codons and regions of the *traD*$_F$ of *R. bellii* may have caused problems at the level of protein translation, which is a common issue when trying to express foreign, AT-rich genes in *E. coli*. Some solutions are to express these genes in a codon-optimized *E. coli* host (Baca and Hol 2000), to add a leader sequence that favors high levels of expression (Ishida and Oshima 2002), to facilitate protein folding by co-expressing a chaperone with the proteins encoded on the expression vector (Liberek *et al.* 1991), and/or by optimizing codons usage of the genes to fit *E. coli*. In this study, a strain of *E. coli* supplemented with rare tRNAs (Rosetta 2(DE3) pLysS) was cultured in M9 minimal media, superior broth, and 2x YT media under kanamycin and chloremphenical selection, but TraD$_F$ was not detected after IPTG induction. An addition of a small open reading frame, or a leader sequence, may be used to allow expression of TraD$_F$ by

affecting the initiation of translation (Ishida and Oshima 2002). This may not be applicable here because our gene was cloned into pet28A that has an N-terminal His-tag that should initiate translation in *E. coli*. Furthermore, TraD was observed controlling for this aspect. Another potential issue is the folding of the protein that may require co-expression of a chaperone protein. However, this is also unlikely to be a problem because the transformants of BL21(DE3) and Rosetta 2(DE3) pLysS were cultured at 15°C overnight to favor a slower and more accurate folding of the protein, yet $TraD_F$ was still not observed. A final option would be to codon-optimize the $traD_F$ gene sequence for expression in *E. coli* using one of several tools available (e.g., Integrated DNA Technology online tool to optimize the codons of a gene to favor expression in *E. coli*) taking care to avoid any secondary RNA structure that may inhibit translation in *E. coli*. If this is successful, then this experimental design of complementing a *traD* mutant with $traD_F$ of *R. bellii* can be executed.

The presence of the *tra* genes of *R. bellii* in other *Rickettsia* species continues to support these genes as a functional unit and not as a degenerative, inactive set of genes. The question at hand is whether these genes have retained their original role in bacterial conjugation or have been adapted for another function. If these genes are indeed involved in rickettsial conjugation, we can conclude that they would have played a major role in the genetic evolution of the genus *Rickettsia* by facilitating acquisition and exchange of genes that enabled adaptation to the changing intracellular environments presented by their disparate hosts. Moreover, an active conjugal system will continue to shape the evolution of rickettsiae.

Fig. 4.1. pet28A-*traD*$_F$ and -*traD* constructs and verification. A) A diagram of a general construct of pet28A-*traD* representing both *traD* of the F-plasmid and *traD*$_F$ of *R. bellii* RML 369-C where EcoRI site was used for amino-terminus of *traD* and *traD*$_F$ and SalI and SacI for the carboxyl-terminus of *traD* and *traD*$_F$. B) The constructs were digested with restriction enzyme EcoRI-SalI and EcoRI-SacI for *traD* of the F-plasmid and *traD*$_F$ of *R. bellii* RML 369-C, respectively, then electrophoresed into a 1% agarose gel and stained with 1x Gel Green. The black asterisk indicates the size of pet28A, white asterisk indicates the size of *traD*$_F$, and black arrow indicates the size of *traD*. Lanes: 1) 1 kb+ ladder; 2-6 and 12) EcoRI-SalI digest of *E. coli* clones transformed with pet28A-*traD*; 5-11) EcoRI-SacI digest of *E. coli* clones transformed with pet28A- *traD*$_F$.

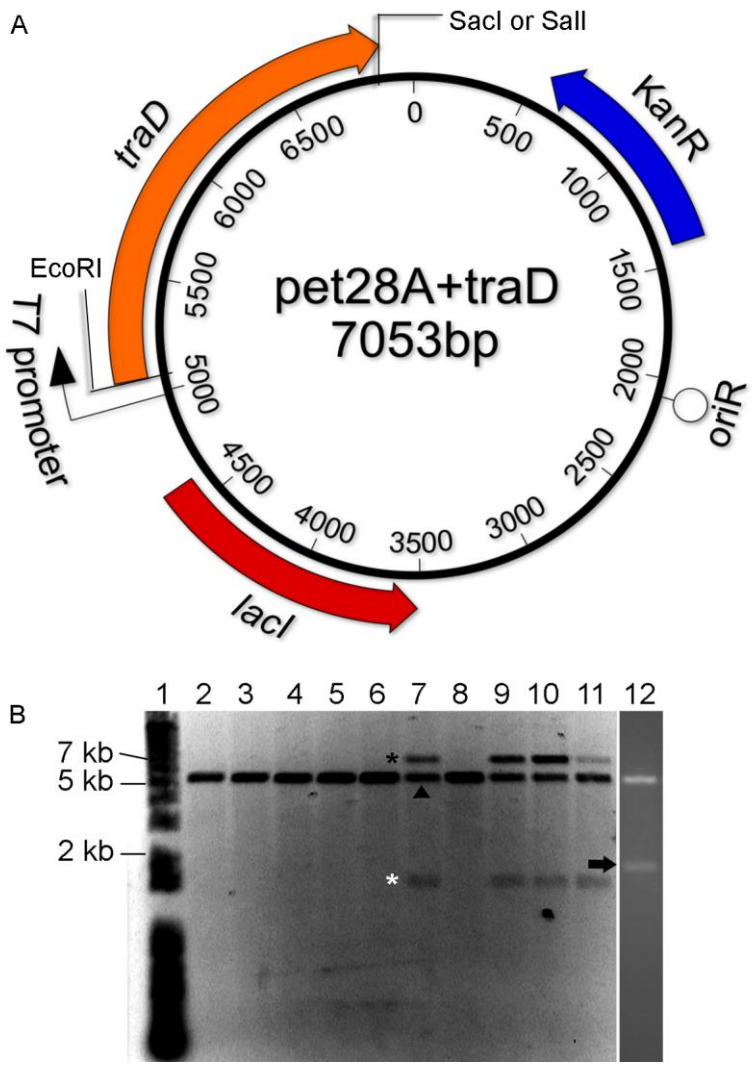

A
SacI or SalI
traD
KanR
EcoRI
T7 promoter
pet28A+traD
7053bp
oriR
lacI
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6500
B
1 2 3 4 5 6 7 8 9 10 11 12
7 kb
5 kb
2 kb

Fig. 4.2. Diagram of plasmids used to construct F+SG and F-*traD36*+SG and verification. A) Diagram of plasmid A representing the wild type F-plasmid and the F-plasmid with *traD36*. Yellow fragment is the *"tra* + oriT/V" fragment from a HindIII digestion to contain the minimal requirement for heritable bacterial conjugation: *tra* cluster, oriT, and oriV (origin of replication). B) Diagram of plasmid B representing pRAM18dSGK#23[MCS] #50. Sky blue fragment is the SG fragment from a HindIII digestion that contains the selection marker (Spec = spectinomycin) and reporter (GFPuv = green fluorescent protein). C) PFGE of HindIII digestion of Stbl4 colonies transformed with the ligation of HindIII-digested plasmid A and the SG fragment from plasmid B. Lanes: 1) 5 kb ladder, 2) 1 kb+ ladder, 3-6) clones of Stbl4 transformed with wild type F-plasmid, 7) blank, 8-11) clones of Stbl4 transformed with F-plasmid with *traD36*. White asterisk indicates size of diagnostic fragments: 55 kb for "*tra* + oriT/V" (yellow fragment of plasmid A) and 3 kb for SG fragment (sky blue fragment of plasmid B). D) PCR detection of *traM* and *traD* of plasmids F+SG (lane 8 in part C) and F-*traD36*+SG (lane 4 in part C) to verify the presence of the *E. coli tra* genes are present in the isolated plasmids. Lanes: 1) 100 bp ladder, *traD*: 2) H_2O, 3) *E. coli* K-12 DNA, 4) F-*traD36*+SG, 5) F+SG, *traM*: 6) H2O, 7) *E. coli* K-12 DNA, 8) F-*traD36*+SG, 9) F+SG.

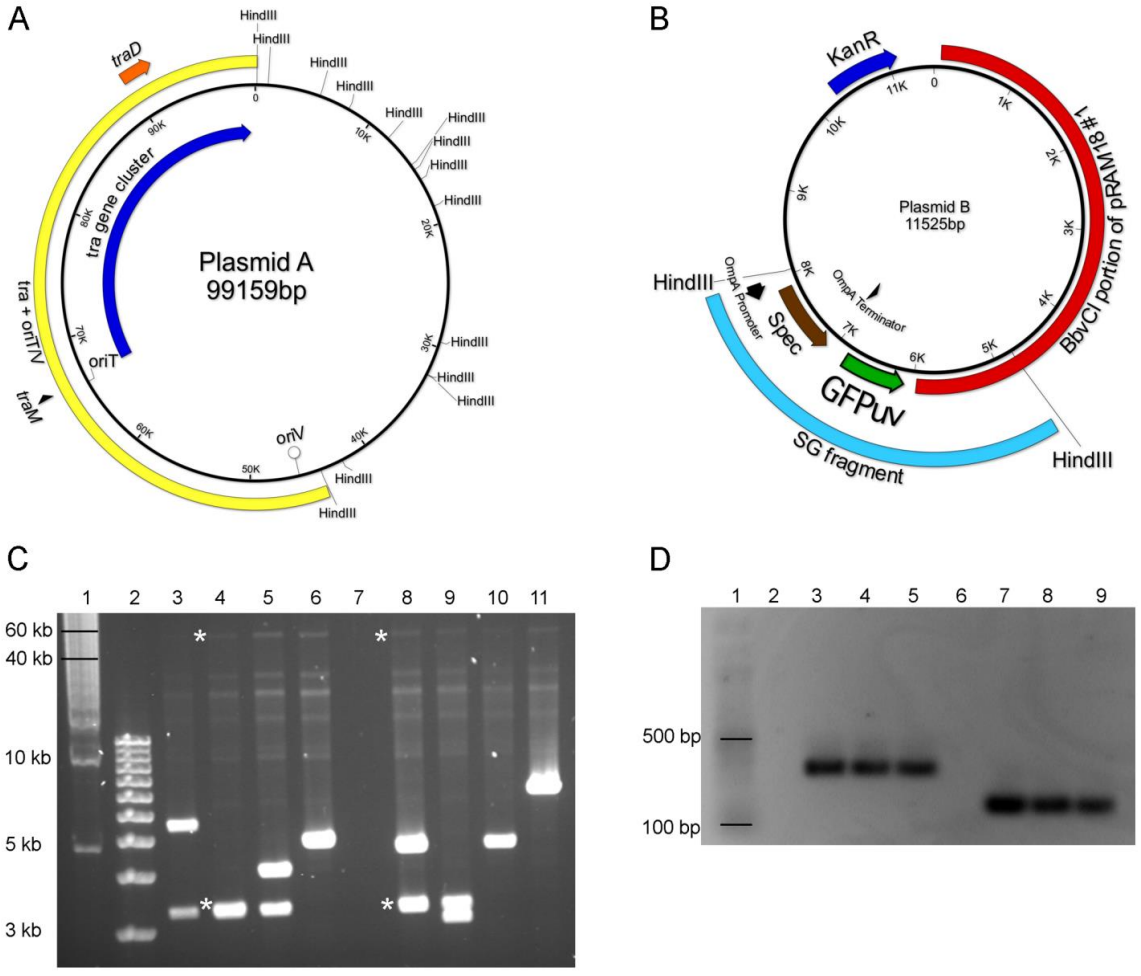

A
Plasmid A
99159bp
tra gene cluster
oriT
oriV
traD
traM
tra + oriTV
HindIII
B
Plasmid B
11525bp
KanR
BbvCI portion of pRAM18#1
HindIII
Omp4 Promoter
Spec
Omp4 Terminator
GFPuv
SG fragment
HindIII
C
1 2 3 4 5 6 7 8 9 10 11
60 kb
40 kb
10 kb
5 kb
3 kb
D
1 2 3 4 5 6 7 8 9
500 bp
100 bp

Fig. 4.3. TraD and $TraD_F$ expression in BL21(DE3). A and B) Coomassie Brilliant Blue (CBB) stained protein gel of BL21(DE3) with pet28A-*traD* or pet28A-$traD_F$, respectively. Lanes are as follows: BL21(DE3) cultured in M9-minimal media (lanes 1-3 in A and 2-4 in B), superior broth (lanes 4-6 in A and 5-7 in B), and 2xYT media (8-10 for both A and B) induced with 0, 0.1, and 0.4 mM IPTG respective to each triplet of lanes within the same medium. A multicolor standard was used to indicate sizes of the proteins (kDa) (lane 7 in A and 1 in B). C and D) Western blot (WB) of the exact duplicate of A and B, respectively, using His-tag primary antibodies (1:2000) and goat anti-mouse secondary antibodies (1:5000). CBB and WB protein gels represent transformants cultured at different temperature (15, 30, and 37^{o}C) with little to no difference in band patterns and detection. Black asterisk (*) indicates predicted size of *E. coli TraD* protein with the His-tag (~85 kDa), and ">" indicates the presence of sequential histidines in *E. coli* proteins that is detected with the His-tag antibody (~34, 26, and 17 kDa). No distinguishable band was present for *R. bellii* $TraD_F$ expression near the predicted 69 kDa area (B and D).

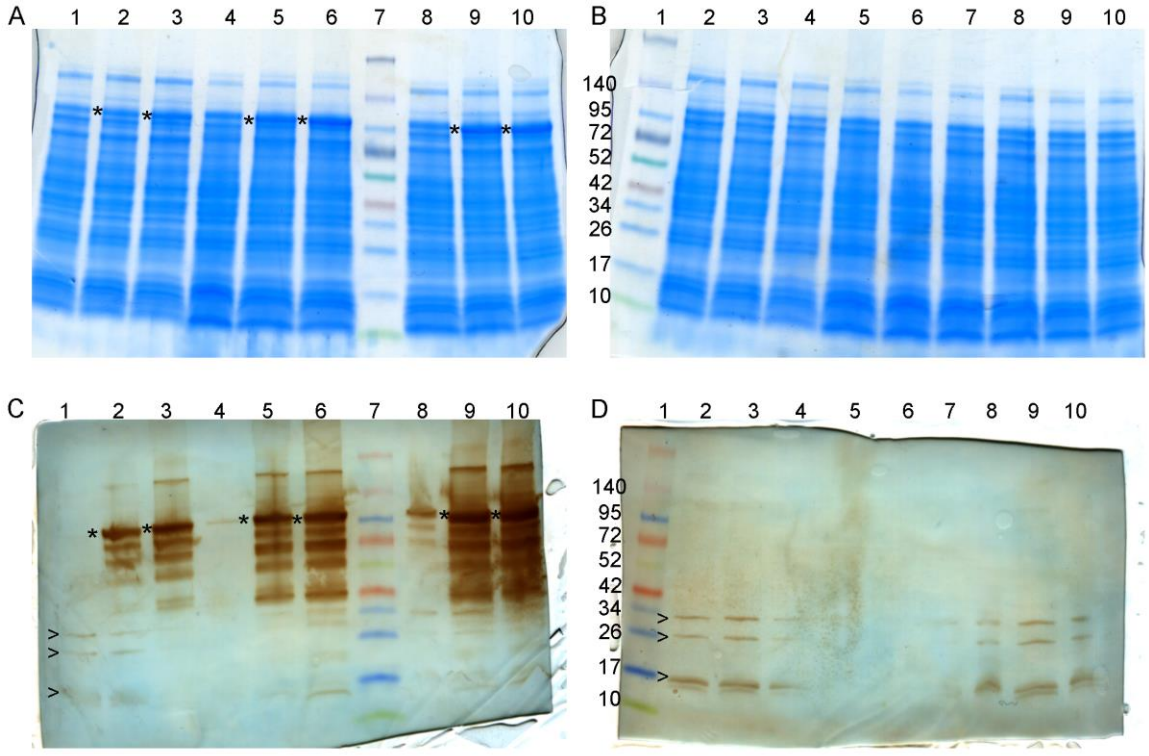

A
1 2 3 4 5 6 7 8 9 10
B
1 2 3 4 5 6 7 8 9 10
140
95
72
52
42
34
26
17
10
C
1 2 3 4 5 6 7 8 9 10
D
1 2 3 4 5 6 7 8 9 10
140
95
72
52
42
34
26
17
10

Fig. 4.4. TraD and $TraD_F$ expression in Rosetta 2(DE3) pLysS. A and B) CBB stained protein gel of Rosetta 2(DE3) pLysS with pet28A-*traD* or pet28A-$traD_F$; respectively, cultured in M9-minimal media (lanes 1-3 in A and 2-4 in B), superior broth (lanes 4-6 in A and 5-7 in B), and 2x YT media (8-10 for both A and B) induced with 0, 0.1, and 0.4 mM IPTG respective to each triplet of lanes. A multicolor standard was used to indicate sizes of the proteins (kDa). C and D) WB of the exact duplicate of A and B, respectively, using His-tag primary antibodies (1:2000) and goat anti-mouse secondary antibodies (1:5000). CBB and WB protein gels represent transformants cultured at different temperature (15, 30, and 37^{o}C) with little to no difference in band patterns and detection. Black asterisk (*) indicates predicted size of *E. coli* TraD protein with the His-tag (~85 kDa). No distinguishable band was present for *R. bellii* $TraD_F$ expression near the predicted 69 kDa area (B and D).

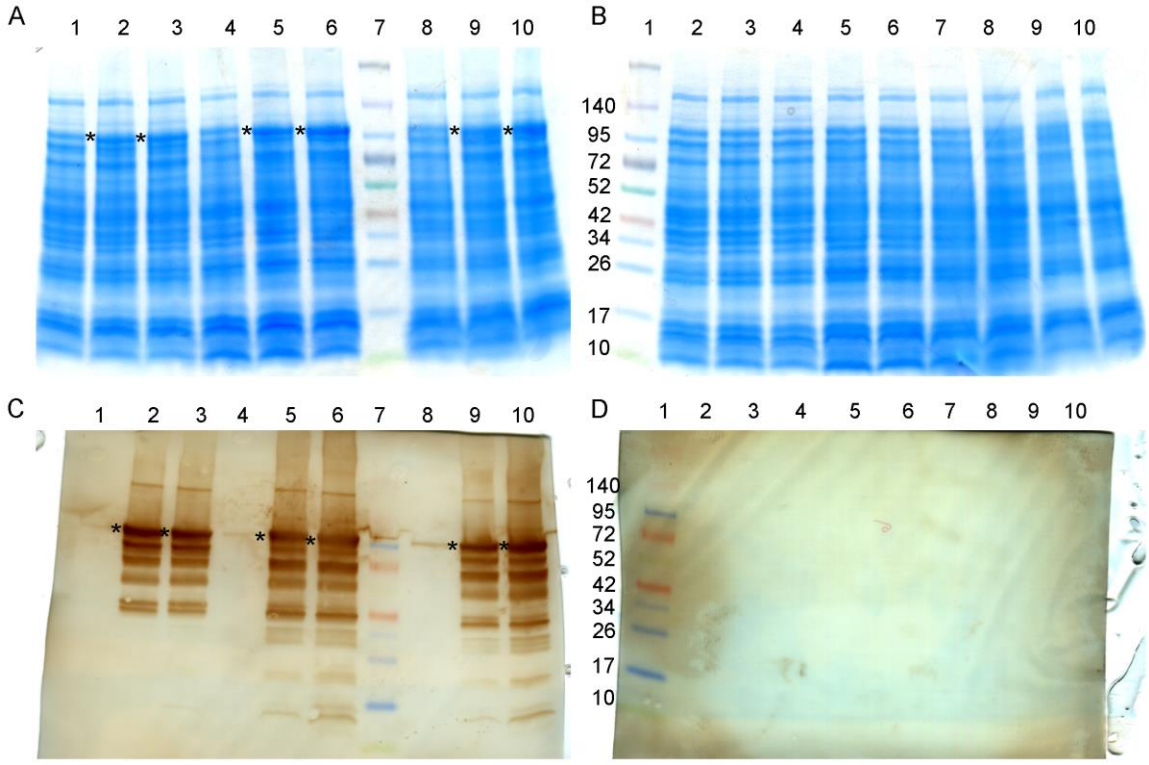

A
1 2 3 4 5 6 7 8 9 10
B
1 2 3 4 5 6 7 8 9 10
140
95
72
52
42
34
26
17
10
C
1 2 3 4 5 6 7 8 9 10
D
1 2 3 4 5 6 7 8 9 10
140
95
72
52
42
34
26
17
10

Table 4.1. Primers for cloning *traD$_F$*/*traD* and checking *tra* genes

Primer	Sequence	Product size (bp)
Rbe traD$_F$ EcoRI clone	ACTGgaattc CAACAAAATCAAAATAATTTTACTCG	1703
Rbe traD$_F$ SacI clone	ACTGgagctc AC AGT TTG TAA ATT CAT TTG TAA TTT AG	
Ecoli traD$_F$ EcoRI clone	ACTG gaattc AGTTTTAACGCAAAGGATATGACC	2168
Ecoli traD$_F$ SalI clone	ACTA gtcgac GA AAT CAT CTC CCG GCT CA	
traM F check	AATCTGACGGCAGTTATTCAG	207
traM R check	AAGATGCGTATGCTCATTCC	
traD F check	CAGACGCTGGTCAGTTATTC	391
traD R check	CCATCTCTTCTTCCGGTTTC	

Chapter 5: Challenges, conclusion and future directions

Rickettsia bellii RML 369-C was chosen as the system to model the transcriptional regulation and activity of the *tra* genes based on the presence of a complete set of *tra* genes in its genome and its naturally wide host range. In addition, it grows robustly for a rickettsia with a doubling time of 8 hours (Heu *et al.* 2015). Other strains of *R. bellii* have a native plasmid that may facilitate DNA transfer (Baldridge *et al.* 2010). Using *R. bellii* RML 369-C, my dissertation explored the transcriptional regulation and the transcriptional activity of the *tra* genes of *R. bellii* RML 369-C with an attempt to complement a *traD* mutant *E. coli* strain with a rickettsial $TraD_F$ protein.

Challenges

Species of *Rickettsia* are obligate, intracellular bacteria presenting difficulty in handling and culturing these organisms. Compared to the model organism *E. coli* where *E. coli* can be cultured in liquid media and on agar plates, host cell cultures are necessary to culture and maintain rickettsiae. Rickettsiae also grows slower with an approximate doubling time of 8 hours (Heu *et al.* 2015), being the fastest growth rate known among rickettsiae, compared to a doubling time of 20-30 minutes for *E. coli*.

My main challenge was to characterize bacterial conjugation in rickettsiae when evidence of genetic transfer has not been documented. This led me to use the well-characterized bacterial conjugation system of *E. coli* for functional assay of the *tra* genes of *R. bellii* in a complementation assay. The challenge was to express an AT-rich gene from *R. bellii* in GC- and AT- balanced genome of *E. coli*. Furthermore, *E. coli* belongs to the gamma-proteobacteria and is distantly related to *R. bellii*, an alpha-

proteobacterium. These features complicated the complementation assay. In addition, the lack of evidence for genetic transfer via bacterial conjugation suggested that the use of site-directed mutagenesis may not result in any measurable trait. Overall, the attempt to provide experimental evidence of the function of the *tra* genes has been met with various challenges that prevented further investigation.

Conclusions and future directions

In Chapter 2, I concluded that the *tra* genes are transcriptionally active in different host cell cultures and at different temperatures. The transcriptional regulation of the *tra* genes, and perhaps most genes, is loosely regulated similarly to *R. prowazekii* (Woodard and Wood 2011). Furthermore, *in silico* predictions group the *tra* genes into multiple operons differentiating it from the *tra* genes of *E. coli*. For future directions and to verify the predictions of the *tra* gene operons, the putative promoters of the *tra* operons should be evaluated for transcription initiation. The putative promoters can be cloned using PCR and the PCR product can be ligated upstream of a reporter gene, e.g. GFPuv, on a shuttle vector with a selectable marker and transformed into *Rickettsia bellii* via electroporation (Rachek *et al.* 1998). A shuttle vector with the *Anaplasma marginale* transcription regulator (AMTR) promoter upstream of the reporter gene would be a positive control for this experimental transformation (Burkhardt *et al.* 2011). Visual observation of *R. bellii* expressing the reporter gene would indicate the presence of a functional *tra* promoter. Southern blot evidence of the plasmid used to transform rickettsiae can corroborate the success of the transformation (Burkhardt *et al.* 2011). Alternatively, the promoter may initiate transcription at specific conditions (i.e. cell density, growth phase) and the

reporter protein may be observed at those conditions (Heu *et al.* 2015). Nonetheless, this experiment would test for the functionality of the putative promoters.

In Chapter 3, I demonstrated that the best reference gene out of the ten selected were *metG* or *metG* and *nrdF*. The same reference gene(s) can be used to standardize future relative transcriptional analysis of rickettsiae with similar treatments or to form a basis of potential reference genes with different treatments. Using the identified reference gene(s), I showed that the transcription of the *tra* genes of *R. bellii* was highly upregulated at 72 HPI relative to 12 HPI only in tick cell culture, indicating that these genes are very active and contribute to the biology of *R. bellii*. Future research should explore the functions of the *tra* genes guided by bioinformatics to obtain detailed documentation of the mechanism of DNA mobility to and from the chromosome and plasmid, and the demonstration of the process of rickettsial conjugation. One potential experiment to test the integration/excision of an integrative, conjugative element is to construct a plasmid with a potential integration site on the plasmid and transform it into *R. bellii*. Integrative, conjugative elements have been implicated to integrate in both plasmids and chromosomes of bacteria for the benefit of dissemination (Burrus and Waldor 2004). In addition, the integrase of the integrative, conjugative element of *R. bellii* has been characterized to target the tRNA-val gene (Gillespie *et al.* 2012). PCR screening of the tRNA-val gene in transformed rickettsia could reveal the mobility of the integrative, conjugative element that contains the *tra* genes. The plasmid could be isolated from the transformant and screened using PCR to test for the integration conjugative element that contains the complete set of *tra* genes into the plasmid. The

results of this would show whether the integrative, conjugative element can mobilize from the chromosome to the plasmid with the putative integration site.

The *tra* genes of *R. bellii* are located in the chromosome complicating the identification of an oriT site, necessary for the recognition of the relaxase and transfer of the *tra* genes. However, recent discoveries identified a complete set of *tra* genes found on plasmids of *R. buchneri* ISO-7 (Gillespie *et al.* 2012) and *R. felis* LSU-Lb (Gillespie *et al.* 2014). By adding a selection and reference gene to these isolated plasmids, one can transform a rickettsia without any *tra* genes with this plasmid and look for any change in morphology with a potential to identify pili-structures. Using the shuttle vectors to transform the previously transformed rickettsia to express a tagged-relaxase will facilitate the identification of the oriT site when using an antibody against the tag in a chromatin-immuno precipitation sequencing assay (Furey 2012). The identification of an oriT site may facilitate the progress in rickettsial bacterial conjugation.

Interestingly, the cloning of the integrative, conjugative element of *R. bellii* into a shuttle vector and transforming it into a rickettsia would be an alternative if the isolation of the native plasmids of *R. buchneri* and *R. felis* proved difficult. Furthermore, the subcloning of the complete set of bacterial conjugation of *E. coli* and transforming it to rickettsiae to screen for any obvious phenotype (i.e. pili, cell-to-cell contact) would be interesting. This experiment has many assumptions with the most important one being the codon bias of GC-rich codons in the *E. coli tra* genes as opposed to the AT-rich codons in the genetic background of rickettsiae. However, the potential to check for pili or other

phenotypes associated with bacterial conjugation may be worth trying and may accelerate research in the field of bacterial conjugation in the species of *Rickettsia*.

In Chapter 4, I attempted to use full length $TraD_F$ of *R. bellii* to rescue genetic transfer in a TraD deficient *E. coli* strain. IPTG induction of BL21(DE3) or Rosetta 2(DE3) pLysS transformed with pet28A-*traD*$_F$ did not show TraD expression despite my testing different temperature, media, IPTG concentration, and codon optimization of *traD*$_F$. For the immediate future direction, a fusion of the *traD* of *E. coli* and *traD*$_F$ of *R. bellii* should be explored instead of attempting to express the full-length, wildtype *traD*$_F$. A two protein hybrid of either the amino-terminus of the TraD or $TraD_F$ fused with the $TraD_F$ or TraD carboxyl-terminus, respectively, should be tested first for expression. The fusion gene could be achieved using PCR cloning of each ends using a common restriction enzyme site to fuse the genes together. The expression of the fusion genes are predicted to express at similar levels as wildtype *traD* of *E. coli*. If successful, these fused TraD proteins may be used in the complementation of a *traD* mutant *E. coli* to rescue genetic transfer. Each fusion protein will indicate the function of half of the *traD*$_F$ of *R. bellii*. Controls would include the expression of only the amino-terminus or the carboxyl-terminus of TraD of *E. coli* in the complementation of the mutant *traD E. coli*. The addition of the amino-terminus of the TraD of *E. coli* may stimulate expression of the carboxyl-terminus of $TraD_F$ of *R. bellii* because it may act as a leader sequence to promote translation of the rest of the fusion protein (Ishida and Oshima 2002). Protein expression can be optimized following this (Sivashanmugam *et al.* 2009). In addition,

$TraA_{Ti}$ of *R. bellii* can be fused with TraI of *E. coli*, similarly as proposed with TraD, to assess the functions of the different parts of $TraA_{Ti}$.

The proposed future experiments will enlighten the function of the Tra proteins as to whether they retained their original function in bacterial conjugation, have been adapted to a different biological function, or are on their way to degeneration. Understanding the rickettsial *tra* genes may further provide deeper insight into the genetic evolution, functional genomics, and the adaptive and changing biology of the genus *Rickettsia*.

References

Adjemian J, Parks S, McElroy K *et al.* Murine typhus in Austin, Texas, USA, 2008. *Emerg Infect Dis* 2010;**16**:412–7.

Alt-Morbe J, Stryker JL, Fuqua C *et al.* The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. *J Bacteriol* 1996;**178**:4248–57.

Alvarez-Martinez CE, Christie PJ. Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* 2009;**73**:775–808.

Andersen CL, Jensen LJ, Ørntoft TF *et al.* Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 2004;**64**:5245–50.

Andersson SGE, Zomorodipour A, Andersson JO *et al.* The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* 1998;**396**:133–40.

Andreotti R, Perez de Leon AA, Dowd SE *et al.* Assessment of bacterial diversity in the cattle tick *Rhipicephalus (Boophilus) microplus* through tag-encoded pyrosequencing. *BMC Microbiol* 2011;**11**:6.

Angelakis E, Botelho E, Socolovschi C *et al.* Murine typhus as a cause of fever in travelers from Tunisia and Mediterranean areas. *J Travel Med* **17**:310–5.

Anthony KG, Klimke WA, Manchak J *et al.* Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *J Bacteriol* 1999;**181**:5149–59.

Apte-Deshpande A, Paingankar M, Gokhale MD *et al. Serratia odorifera* a midgut inhabitant of Aedes aegypti mosquito enhances its susceptibility to dengue-2 virus. *PLoS One* 2012;**7**:e40401.

Arechaga I, Pena A, Zunzunegui S *et al.* ATPase activity and oligomeric state of TrwK, the VirB4 homologue of the plasmid R388 type IV secretion system. *J Bacteriol* 2008;**190**:5472–9.

Arutyunov D, Arenson B, Manchak J *et al.* F plasmid TraF and TraH are components of an outer membrane complex involved in conjugation. *J Bacteriol* 2010;**192**:1730–4.

Audette GF, Manchak J, Beatty P *et al.* Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology* 2007;**153**:442–51.

Azad AF. Epidemiology of murine typhus. *Annu Rev Entomol* 1990;**35**:553–70.

Baca AM, Hol WGJ. Overcoming codon bias: A method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli. *Int J Parasitol* 2000;**30**:113–8.

Baldridge GD, Burkhardt NY, Felsheim RF *et al.* Transposon insertion reveals pRM, a plasmid of *Rickettsia monacensis*. *Appl Environ Microbiol* 2007a;**73**:4984–95.

Baldridge GD, Burkhardt NY, Felsheim RF *et al.* Plasmids of the pRM/pRF family occur in diverse *Rickettsia* species. *Appl Environ Microbiol* 2008;**74**:645–52.

Baldridge GD, Burkhardt NY, Labruna MB *et al.* Wide dispersal and possible multiple

origins of low-copy-number plasmids in rickettsia species associated with blood-feeding arthropods. *Appl Environ Microbiol* 2010;**76**:1718–31.
Baldridge GD, Kurtti TJ, Burkhardt N *et al.* Infection of *Ixodes scapularis* ticks with *Rickettsia monacensis* expressing green fluorescent protein: a model system. *J Invertebr Pathol* 2007b;**94**:163–74.
Barbieri AR, Romero L, Labruna MB. *Rickettsia bellii* infecting *Amblyomma sabanerae* ticks in El Salvador. *Pathog Glob Health* 2012;**106**:188–9.
de Barros Lopes L, Guterres A, Rozental T *et al. Rickettsia bellii, Rickettsia amblyommii*, and Laguna Negra hantavirus in an Indian reserve in the Brazilian Amazon. *Parasit Vectors* 2014;**7**:191.
Bechah Y, Capo C, Mege J-L *et al.* Epidemic typhus. *Lancet Infect Dis* 2008;**8**:417–26.
Bellanger X, Payot S, Leblond-Bourget N *et al.* Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiol Rev* 2014;**38**:720–60.
Bendtsen JD, Nielsen H, von Heijne G *et al.* Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;**340**:783–95.
Blanc G, Ogata H, Robert C *et al.* Lateral gene transfer between obligate intracellular bacteria: evidence from the *Rickettsia massiliae* genome. *Genome Res* 2007a;**17**:1657–64.
Blanc G, Ogata H, Robert C *et al.* Reductive genome evolution from the mother of *Rickettsia*. *PLoS Genet* 2007b;**3**:e14.
Brady CM, Asplen MK, Desneux N *et al.* Worldwide populations of the aphid *Aphis craccivora* are infected with diverse facultative bacterial symbionts. *Microb Ecol* 2014;**67**:195–204.
Budachetri K, Browning RE, Adamson SW *et al.* An insight into the microbiome of the *Amblyomma maculatum* (Acari: Ixodidae). *J Med Entomol* 2014;**51**:119–29.
Burkhardt NY, Baldridge GD, Williamson PC *et al.* Development of shuttle vectors for transformation of diverse *Rickettsia* species. *PLoS One* 2011;**6**:e29511.
Burrus V, Waldor MK. Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* 2004;**155**:376–86.
Cabezon E, Lanka E, de la Cruz F. Requirements for mobilization of plasmids RSF1010 and ColE1 by the IncW plasmid R388: trwB and RP4 traG are interchangeable. *J Bacteriol* 1994;**176**:4455–8.
Cabezon E, Ripoll-Rozada J, Pena A *et al.* Towards an integrated model of bacterial conjugation. *FEMS Microbiol Rev* 2015;**39**:81–95.
Campbell JL, Richardson CC, Studier FW. Genetic recombination and complementation between bacteriophage T7 and cloned fragments of T7 DNA. *Proc Natl Acad Sci U S A* 1978;**75**:2276–80.
Cardwell MM, Martinez JJ. The Sca2 autotransporter protein from *Rickettsia conorii* is sufficient to mediate adherence to and invasion of cultured mammalian cells. *Infect Immun* 2009;**77**:5272–80.
Carmichael JR, Fuerst PA. Molecular Detection of *Rickettsia bellii , Rickettsia montanensis ,* and *Rickettsia rickettsii* in a *Dermacentor variabilis* Tick from Nature. *Vector-Borne Zoonotic Dis* 2010;**10**:111–5.
Caspi-Fluger A, Inbar M, Steinberg S *et al.* Characterization of the symbiont *Rickettsia* in

the mirid bug *Nesidiocoris tenuis* (Reuter) (Heteroptera: Miridae). *Bull Entomol Res* 2014;**104**:681–8.
Cho BK, Barrett CL, Knight EM *et al.* Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc Natl Acad Sci U S A* 2008;**105**:19462–7.
Cho H, Winans SC. TraA, TraC and TraD autorepress two divergent quorum-regulated promoters near the transfer origin of the Ti plasmid of *Agrobacterium tumefaciens*. *Mol Microbiol* 2007;**63**:1769–82.
Choi YJ, Jang WJ, Kim JH *et al.* Spotted fever group and typhus group rickettsioses in humans, South Korea. *Emerg Infect Dis* 2005;**11**:237–44.
Civen R, Ngo V. Murine typhus: an unrecognized suburban vectorborne disease. *Clin Infect Dis* 2008;**46**:913–8.
Clay K, Klyachko O, Grindle N *et al.* Microbial communities and interactions in the lone star tick, *Amblyomma americanum*. *Mol Ecol* 2008;**17**:4371–81.
Dam P, Olman V, Harris K *et al.* Operon prediction using both genome-specific and general genomic information. *Nucleic Acids Res* 2007;**35**:288–98.
Doran TJ, Loh SM, Firth N *et al.* Molecular analysis of the F plasmid *traVR* region: *traV* encodes a lipoprotein. *J Bacteriol* 1994;**176**:4182–6.
Dumler JS. *Intracellular Pathogens II: Rickettsiales*. Azad AF, Palmer GH (eds.). American Society of Microbiology, 2012.
Duron O. Lateral transfers of insertion sequences between *Wolbachia*, *Cardinium* and *Rickettsia* bacterial endosymbionts. *Heredity (Edinb)* 2013;**111**:330–7.
Ellison DW, Clark TR, Sturdevant DE *et al.* Genomic comparison of virulent *Rickettsia rickettsii* Sheila Smith and avirulent *Rickettsia rickettsii* Iowa. *Infect Immun* 2008;**76**:542–50.
Emelyanov V V. Rickettsiaceae, rickettsia-like endosymbionts, and the origin of mitochondria. *Biosci Rep* 2001;**21**:1–17.
Eremeeva ME, Dasch GA, Silverman DJ. Evaluation of a PCR assay for quantitation of *Rickettsia rickettsii* and closely related spotted fever group rickettsiae. *J Clin Microbiol* 2003;**41**:5466–72.
Eremeeva ME, Karpathy SE, Levin ML *et al.* Spotted fever rickettsiae, *Ehrlichia* and *Anaplasma*, in ticks from peridomestic environments in Panama. *Clin Microbiol Infect* 2009;**15 Suppl 2**:12–4.
Eremeeva ME, Madan A, Shaw CD *et al.* New perspectives on rickettsial evolution from new genome sequences of rickettsia, particularly *R. canadensis*, and *Orientia tsutsugamushi*. *Ann N Y Acad Sci* 2005;**1063**:47–63.
Felsheim RF, Kurtti TJ, Munderloh UG. Genome sequence of the endosymbiont *Rickettsia peacockii* and comparison with virulent *Rickettsia rickettsii*: identification of virulence factors. *PLoS One* 2009;**4**:e8361.
Firth N, Ippen-Ihler K, Skurray RA. Structure and Function of the F Factor and Mechanism of Conjugation. In: Neidhardt FC (ed.). *In Escherichia Coli and Salmonella: Cellular and Molecular Biology*. Washington, DC: American Society for Microbiology, 1996, 2377–401.
Firth N, Skurray R. Characterization of the F plasmid bifunctional conjugation gene, *traG*. *Mol Gen Genet* 1992;**232**:145–53.

Francia M V, Varsaki A, Garcillan-Barcia MP *et al.* A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol Rev* 2004;**28**:79–100.
Francis F, Guillonneau F, Leprince P *et al.* Tritrophic interactions among *Macrosiphum euphorbiae* aphids, their host plants and endosymbionts: investigation by a proteomic approach. *J Insect Physiol* 2010;**56**:575–85.
Frost LS, Ippen-Ihler K, Skurray RA. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* 1994;**58**:162–210.
Frost LS, Manchak J. F- phenocopies: characterization of expression of the F transfer region in stationary phase. *Microbiology* 1998;**144**:2579–87.
Frost LS, Paranchych W, Willetts NS. DNA sequence of the F *traALE* region that includes the gene for F pilin. *J Bacteriol* 1984;**160**:395–401.
Furey TS. ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. *Nat Rev Genet* 2012;**13**:840–52.
Fuxelius H-HH, Darby AC, Cho N-HH *et al.* Visualization of pseudogenes in intracellular bacteria reveals the different tracks to gene destruction. *Genome Biol* 2008;**9**:R42–2008 – 9–2 – r42. Epub 2008 Feb 26.
Galletti MF, Fujita A, Nishiyama Jr MY *et al.* Natural blood feeding and temperature shift modulate the global transcriptional profile of *Rickettsia rickettsii* infecting its tick vector. *PLoS One* 2013;**8**:e77388.
Garcillan-Barcia MP, Francia M V, de la Cruz F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* 2009;**33**:657–87.
Gaywee J, Radulovic S, Higgins JA *et al.* Transcriptional analysis of *Rickettsia prowazekii* invasion gene homolog (invA) during host cell infection. *Infect Immun* 2002;**70**:6346–54.
Gebiola M, Gomez-Zurita J, Monti MM *et al.* Integration of molecular, ecological, morphological and endosymbiont data for species delimitation within the *Pnigalio soemius* complex (Hymenoptera: Eulophidae). *Mol Ecol* 2012;**21**:1190–208.
Georgiades K, Merhej V, El Karkouri K *et al.* Gene gain and loss events in *Rickettsia* and *Orientia* species. *Biol Direct* 2011;**6**:6.
Gerth M, Saeed A, White JA *et al.* Extensive screen for bacterial endosymbionts reveals taxon-specific distribution patterns among bees (Hymenoptera, Anthophila). *FEMS Microbiol Ecol* 2015;**91**:10.1093/femsec/fiv047. Epub 2015 Apr 26.
Gillespie JJ, Beier MS, Rahman MS *et al.* Plasmids and rickettsial evolution: insight from *Rickettsia felis*. *PLoS One* 2007;**2**:e266.
Gillespie JJ, Brayton K a., Williams KP *et al.* Phylogenomics reveals a diverse *Rickettsiales* type IV secretion system. *Infect Immun* 2010;**78**:1809–23.
Gillespie JJ, Driscoll TP, Verhoeve VI *et al.* Genomic diversification in strains of *Rickettsia felis Isolated* from different arthropods. *Genome Biol Evol* 2014;**7**:35–56.
Gillespie JJ, Joardar V, Williams KP *et al.* A *Rickettsia* genome overrun by mobile genetic elements provides insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. *J Bacteriol* 2012;**194**:376–94.
Gillespie JJ, Kaur SJ, Rahman MS *et al.* Secretome of obligate intracellular *Rickettsia*. *FEMS Microbiol Rev* 2015;**39**:47–80.
Gillespie JJ, Williams K, Shukla M *et al. Rickettsia* phylogenomics: unwinding the

intricacies of obligate intracellular life. *PLoS One* 2008;**3**:e2018.
Gillespie JJJ, Ammerman NCNNC, Dreher-Lesnick SMS *et al.* An anomalous type IV secretion system in *Rickettsia* is evolutionarily conserved. *PLoS One* 2009;**4**:e4833.
Giorgini M, Bernardo U, Monti MM *et al. Rickettsia* symbionts cause parthenogenetic reproduction in the parasitoid wasp *Pnigalio soemius* (Hymenoptera: Eulophidae). *Appl Environ Microbiol* 2010;**76**:2589–99.
Gomis-Ruth FX, Sola M, de la Cruz F *et al.* Coupling factors in macromolecular type-IV secretion machineries. *Curr Pharm Des* 2004;**10**:1551–65.
Gottlieb Y, Ghanim M, Chiel E *et al.* Identification and localization of a Rickettsia sp. in Bemisia tabaci (Homoptera: Aleyrodidae). *Appl Environ Microbiol* 2006;**72**:3646–52.
Gouin E, Egile C, Dehoux P *et al.* The RickA protein of *Rickettsia conorii* activates the Arp2/3 complex. *Nature* 2004;**427**:457–61.
Guglielmini J, Neron B, Abby SS *et al.* Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res* 2014;**42**:5715–27.
Haglund CM, Choe JE, Skau CT *et al. Rickettsia* Sca2 is a bacterial formin-like mediator of actin-based motility. *Nat Cell Biol* 2010;**12**:1057–63.
Haglund CM. *Rickettsia* Sca2 is a bacterial formin-like mediator of actin-based motility. *Univ Calif Berkeley* 2011;**Ph.D.**
Hamilton CM, Lee H, Li PL *et al.* TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J Bacteriol* 2000;**182**:1541–8.
Harden VA. *Rocky Mountain Spotted Fever: History of a Twentieth-Century Disease*. Baltimore and London: The John Hopkins University Press, 1990.
Harris RL, Hombs V, Silverman PM. Evidence that F-plasmid proteins TraV, TraK and TraB assemble into an envelope-spanning structure in *Escherichia coli*. *Mol Microbiol* 2001;**42**:757–66.
Harris RL, Silverman PM. Roles of internal cysteines in the function, localization, and reactivity of the TraV outer membrane lipoprotein encoded by the F plasmid. *J Bacteriol* 2002;**184**:3126–9.
Harris RL, Silverman PM. Tra proteins characteristic of F-like type IV secretion systems constitute an interaction group by yeast two-hybrid analysis. *J Bacteriol* 2004;**186**:5480–5.
Hayat S, Elofsson A. BOCTOPUS: improved topology prediction of transmembrane beta barrel proteins. *Bioinformatics* 2012;**28**:516–22.
Hellemans J, Mortier G, De Paepe A *et al.* qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 2007;**8**:R19.
Hendry TA, Hunter MS, Baltrus DA. The facultative symbiont *Rickettsia* protects an invasive whitefly against entomopathogenic *Pseudomonas syringae* strains. *Appl Environ Microbiol* 2014;**80**:7161–8.
Hernandez-Lopez A, Chabrol O, Royer-Carenzi M *et al.* To tree or not to tree? Genome-wide quantification of recombination and reticulate evolution during the

diversification of strict intracellular bacteria. *Genome Biol Evol* 2013;**5**:2305–17.
Heu CC, Kurtti TJ, Nelson CM *et al.* Transcriptional analysis of the conjugal transfer tenes of *Rickettsia bellii* RML 369-C. *PLoS One* 2015;**10**:e0137214.
Hopps HE, Jackson EB, Danauskas JX *et al.* Study on the growth of rickettsiae. III. Influence of extracellular environment on the growth of *Rickettsia tsutsugamushi* in tissue culture cells. *J Immunol (Baltimore, Md 1950)* 1959;**82**:161–71.
Horta MC, Pinter A, Schumaker TTS *et al.* Natural infection, transovarial transmission, and transstadial survival of *Rickettsia bellii* in the tick *Ixodes loricatus* (Acari: Ixodidae) from Brazil. *Ann N Y Acad Sci* 2006;**1078**:285–90.
Howard MT, Nelson WC, Matson SW. Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem* 1995;**270**:28381–6.
Hoy MA, Jeyaprakash A. Microbial diversity in the predatory mite *Metaseiulus occidentalis* (Acari: Phytoseiidae) and its prey, *Tetranychus urticae* (Acari: Tetranychidae). *Biol Control* 2005;**32**:427–41.
Ishida M, Oshima T. Effective structure of a leader open reading frame for enhancing the expression of GC-rich genes. *J Biochem* 2002;**132**:63–70.
Juncker AS, Willenbrock H, Von Heijne G *et al.* Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* 2003;**12**:1652–62.
Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* 2004;**338**:1027–36.
Kall L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res* 2007;**35**:W429–32.
Klimke WA, Rypien CD, Klinger B *et al.* The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. *Microbiology* 2005;**151**:3527–40.
Koga R, Bennett GM, Cryan JR *et al.* Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. *Environ Microbiol* 2013;**15**:2073–81.
Koraimann G, Teferle K, Mitteregger R *et al.* Differential mRNA decay within the transfer operon of plasmid R1: identification and analysis of an intracistronic mRNA stabilizer. *Mol Gen Genet* 1996;**250**:466–76.
Kristoffersen SM, Haase C, Weil MR *et al.* Global mRNA decay analysis at single nucleotide resolution reveals segmental and positional degradation patterns in a Gram-positive bacterium. *Genome Biol* 2012;**13**:R30–2012 – 13–4 – r30.
Krogh A, Larsson B, von Heijne G *et al.* Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;**305**:567–80.
Kurtti TJ, Felsheim RF, Burkhardt NY *et al. Rickettsia buchneri* sp. nov., a rickettsial endosymbiont of the blacklegged tick *Ixodes scapularis*. *Int J Syst Evol Microbiol* 2015;**65**:965–70.
Labruna MB, Pacheco RC, Nava S *et al.* Infection by *Rickettsia bellii* and Candidatus "Rickettsia amblyommii" in *Amblyomma neumanni* ticks from Argentina." *Microl Ecol* 2007a;**54**:126–33.
Labruna MB, Pacheco RC, Richtzenhain LJ *et al.* Isolation of *Rickettsia rhipicephali* and

Rickettsia bellii from *Haemaphysalis juxtakochi* ticks in the State of São Paulo, Brazil. *Appl Environ Microbiol* 2007b;**73**:869–73.
Labruna MB, Whitworth T, Bouyer DH *et al. Rickettsia bellii* and *Rickettsia amblyommii* in *Amblyomma* ticks from the State of Rondonia, Western Amazon, Brazil. *J Med Entomol* 2004a;**41**:1073–81.
Labruna MB, Whitworth T, Horta MC *et al. Rickettsia* species infecting *Amblyomma cooperi* ticks from an area in the state of Sao Paulo, Brazil, where Brazilian spotted fever is endemic. *J Clin Microbiol* 2004b;**42**:90–8.
Lalzar I, Harrus S, Mumcuoglu KY *et al.* Composition and seasonal variation of *Rhipicephalus turanicus* and *Rhipicephalus sanguineus* bacterial communities. *Appl Environ Microbiol* 2012;**78**:4110–6.
Lang S, Kirchberger PC, Gruber CJ *et al.* An activation domain of plasmid R1 TraI protein delineates stages of gene transfer initiation. *Mol Microbiol* 2011;**82**:1071–85.
Lawley TD, Klimke WA, Gubbins MJ *et al.* F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* 2003;**224**:1–15.
Lawrence A-M, Besir HUS. Staining of proteins in gels with Coomassie G-250 without organic solvent and acetic acid. *J Vis Exp* 2009:2–4.
Lawson ET, Mousseau T a., Klaper R *et al. Rickettsia* associated with male-killing in a buprestid beetle. *Heredity (Edinb)* 2001;**86**:497–505.
Liberek K, Marszalek J, Ang D *et al.* Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc Natl Acad Sci U S A* 1991;**88**:2874–8.
Liu D, Reeves PR. *Escherichia coli* K12 regains its O-antigen. *Microbiology-Uk* 1994;**140**:49–57.
Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods* 2001;**25**:402–8.
Llosa M, Zunzunegui S, de la Cruz F. Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc Natl Acad Sci U S A* 2003;**100**:10465–70.
Lommano E, Bertaiola L, Dupasquier C *et al.* Infections and coinfections of questing *Ixodes ricinus* ticks by emerging zoonotic pathogens in Western Switzerland. *Appl Environ Microbiol* 2012;**78**:4606–12.
Lu F, Kang X, Lorenz G *et al.* Culture-independent analysis of bacterial communities in the gut of rice water weevil (Coleoptera: Curculionidae). *Ann Entomol Soc Am* 2014;**107**:592–600.
Machtelinckx T, Van Leeuwen T, Van De Wiele T *et al.* Microbial community of predatory bugs of the genus *Macrolophus* (Hemiptera: Miridae). *BMC Microbiol* 2012;**12 Suppl 1**:S9–2180 – 12–S1 – S9.
Malek JA, Wierzbowski JM, Tao W *et al.* Protein interaction mapping on a functional shotgun sequence of *Rickettsia sibirica*. *Nucleic Acids Res* 2004;**32**:1059–64.
Manea SJ, Sasaki DM, Ikeda JK *et al.* Clinical and epidemiological observations regarding the 1998 Kauai murine typhus outbreak. *Hawaii Med J* 2001;**60**:7–11.
Maneewannakul S, Kathir P, Ippen-Ihler K. Characterization of the F plasmid mating

aggregation gene *traN* and of a new F transfer region locus *trbE*. *J Mol Biol* 1992;**225**:299–311.
Maneewannakul S, Maneewannakul K, Ippen-Ihler K. Characterization, localization, and sequence of F transfer region products: the pilus assembly gene product TraW and a new product, TrbI. *J Bacteriol* 1992;**174**:5567–74.
Manning PA, Morelli G, Achtman M. traG protein of the F sex factor of *Escherichia coli* K-12 and its role in conjugation. *Proc Natl Acad Sci U S A* 1981;**78**:7487–91.
Mao F, Dam P, Chou J *et al.* DOOR: a database for prokaryotic operons. *Nucleic Acids Res* 2009;**37**:D459–63.
Mao X, Ma Q, Zhou C *et al.* DOOR 2.0: presenting operons and their functions through dynamic and integrated views. *Nucleic Acids Res* 2014;**42**:D654–9.
Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 2004;**32**:W327–31.
Marrero J, Waldor MK. Interactions between inner membrane proteins in donor and recipient cells limit conjugal DNA transfer. *Dev Cell* 2005;**8**:963–70.
Marrero J, Waldor MK. Determinants of entry exclusion within Eex and TraG are cytoplasmic. *J Bacteriol* 2007;**189**:6469–73.
Martin OY, Puniamoorthy N, Gubler A *et al.* Infections with Wolbachia, Spiroplasma, and Rickettsia in the Dolichopodidae and other Empidoidea. *Infect Genet Evol* 2013;**13**:317–30.
Matson SW, Ragonese H. The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. *J Bacteriol* 2005;**187**:697–706.
Matson SW, Sampson JK, Byrd DR. F plasmid conjugative DNA transfer: the TraI helicase activity is essential for DNA strand transfer. *J Biol Chem* 2001;**276**:2372–9.
Matsuura Y, Kikuchi Y, Meng XY *et al.* Novel clade of alphaproteobacterial endosymbionts associated with stinkbugs and other arthropods. *Appl Environ Microbiol* 2012;**78**:4149–56.
McIntosh D, Bezerra RA, Luz HR *et al.* Detection of *Rickettsia bellii* and *Rickettsia amblyommii* in *Amblyomma longirostre* (Acari: Ixodidae) from Bahia state, Northeast Brazil. *Braz J Microbiol* 2015;**46**:879–83.
Mcleod MP, Qin X, Karpathy SE *et al.* Complete genome sequence of *Rickettsia typhi* and comparison with sequences of other rickettsiae. *Society* 2004;**186**:5842–55.
Merhej V, Notredame C, Royer-Carenzi M *et al.* The rhizome of life: the sympatric *Rickettsia felis* paradigm demonstrates the random transfer of DNA sequences. *Mol Biol Evol* 2011;**28**:3213–23.
Milhano N, de Carvalho IL, Alves AS *et al.* Coinfections of *Rickettsia slovaca* and *Rickettsia helvetica* with *Borrelia lusitaniae* in ticks collected in a Safari Park, Portugal. *Ticks Tick Borne Dis* 2010;**1**:172–7.
Moliner C, Fournier PE, Raoult D. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. *FEMS Microbiol Rev* 2010;**34**:281–94.
Moncalian G, Cabezon E, Alkorta I *et al.* Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. *J Biol Chem* 1999;**274**:36117–24.
Moreno CX, Moy F, Daniels TJ *et al.* Molecular analysis of microbial communities

identified in different developmental stages of *Ixodes scapularis* ticks from Westchester and Dutchess Counties, New York. *Environ Microbiol* 2006;**8**:761–72.
Munderloh UG, Jauron SD, Fingerle V *et al.* Invasion and intracellular development of the human granulocytic ehrlichiosis agent in tick cell culture. *J Clin Microbiol* 1999;**37**:2518–24.
Munderloh UG, Kurtti TJ. Formulation of medium for tick cell culture. *Exp Appl Acarol* 1989;**7**:219–29.
Murray ES, Baehr G, Shwartzman G *et al.* Brill's Disease: I. Clinical and laboratory diagnosis. *J Am Med Assoc* 1950;**142**:1059.
Nakayama H, Kurokawa K, Lee BL. Lipoproteins in bacteria: structures and biosynthetic pathways. *FEBS J* 2012;**279**:4247–68.
Narasimhan S, Fikrig E. Tick microbiome: the force within. *Trends Parasitol* 2015;**31**:315–23.
Narasimhan S, Rajeevan N, Liu L *et al.* Gut microbiota of the tick vector *Ixodes scapularis* modulate colonization of the Lyme disease spirochete. *Cell Host Microbe* 2014;**15**:58–71.
Nelson WC, Howard MT, Sherman JA *et al.* The traY gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid oriT. *J Biol Chem* 1995;**270**:28374–80.
Niebylski ML, Peacock MG, Fischer ER *et al.* Characterization of an endosymbiont infecting wood ticks, *Dermacentor andersoni*, as a member of the genus *Francisella*. *Appl Environ Microbiol* 1997a;**63**:3933–40.
Niebylski ML, Peacock MG, Schwan TG. Lethal effect of *Rickettsia rickettsii* on its tick vector (*Dermacentor andersoni*). *Appl Environ Microbiol* 1999;**65**:773–8.
Niebylski ML, Schrumpf ME, Burgdorfer W *et al. Rickettsia peacockii* sp. nov., a new species infecting wood ticks, *Dermacentor andersoni*, in western Montana. *Int J Syst Bacteriol* 1997b;**47**:446–52.
Nikoh N, Hosokawa T, Moriyama M *et al.* Evolutionary origin of insect-*Wolbachia* nutritional mutualism. *Proc Natl Acad Sci U S A* 2014;**111**:10257–62.
Ogata H, Audic S, Renesto-Audiffren P *et al.* Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science (80-)* 2001;**293**:2093–8.
Ogata H, Renesto P, Audic S *et al.* The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. *PLoS Biol* 2005a;**3**:e248.
Ogata H, Robert C, Audic S *et al. Rickettsia felis*, from culture to genome sequencing. *Ann N Y Acad Sci* 2005b;**1063**:26–34.
Ogata H, La Scola B, Audic S *et al.* Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet* 2006;**2**:e76.
Ogrzewalska M, Literak I, Cardenas-Callirgos JM *et al. Rickettsia bellii* in ticks *Amblyomma varium* Koch, 1844, from birds in Peru. *Ticks Tick Borne Dis* 2012;**3**:254–6.
Ogrzewalska M, Pacheco RC, Uezu A *et al.* Rickettsial infection in *Amblyomma nodosum* ticks (Acari: Ixodidae) from Brazil. *Ann Trop Med Parasitol*

2009;**103**:413–25.
Ogrzewalska M, Uezu A, Jenkins CN *et al.* Effect of forest fragmentation on tick infestations of birds and tick infection rates by *Rickettsia* in the Atlantic forest of Brazil. *Ecohealth* 2011;**8**:320–31.
Oliver JD, Burkhardt NY, Felsheim RF *et al.* Motility characteristics are altered for *Rickettsia bellii* transformed to overexpress a heterologous *rickA* gene. *Appl Environ Microbiol* 2013, DOI: 10.1128/AEM.03352-13.
Pacheco RC, Horta MC, Pinter A *et al.* Survey of *Rickettsia* spp in the ticks *Amblyomma cajennense* and *Amblyomma dubitatum* in the State of Sao Paulo. *Rev Soc Bras Med Trop* 2009;**42**:351–3.
Pacheco RC, Rosa S, Richtzenhain LJ *et al.* Isolation of *Rickettsia bellii* from *Amblyomma ovale* and *Amblyomma incisum* ticks from southern Brazil. *Rev MVZ Cordoba* 2008;**13**:1273–9.
Paddock CD, Koss T, Eremeeva ME *et al.* Isolation of *Rickettsia akari* from eschars of patients with rickettsialpox. *Am J Trop Med Hyg* 2006;**75**:732–8.
Paddock CD, Sumner JW, Comer JA *et al.* Rickettsia parkeri: a newly recognized cause of spotted fever rickettsiosis in the United States. *Clin Infect Dis* 2004;**38**:805–11.
Panicker MM, Minkley Jr EG. Purification and properties of the F sex factor TraD protein, an inner membrane conjugal transfer protein. *J Biol Chem* 1992;**267**:12761–6.
Parola P, Paddock CD, Socolovschi C *et al.* Update on tick-borne rickettsioses around the world: a geographic approach. *Clin Microbiol Rev* 2013;**26**:657–702.
Parola P. *Rickettsia felis*: From a rare disease in the USA to a common cause of fever in sub-Saharan Africa. *Clin Microbiol Infect* 2011;**17**:996–1000.
Patterson KD. Typhus and its control in Russia, 1870–1940. *Med Hist* 2012;**37**:361–81.
Pfaffl MW, Tichopad A, Prgomet C *et al.* Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett* 2004;**26**:509–15.
Philip RN, Casper E a., Anacker RL *et al. Rickettsia bellii* sp. nov.: a tick-borne *Rickettsia*, widely distributed in the United States, that is distinct from the spotted fever and typhus biogroups. *Int J Syst Bacteriol* 1983;**33**:94–106.
Philip RN, Casper EA. Serotypes of spotted fever group rickettsiae isolated from *Dermacentor andersoni* (Stiles) ticks in western Montana. *Am J Trop Med Hyg* 1981;**30**:230–8.
Pinter A, Labruna MB. Isolation of *Rickettsia rickettsii* and *Rickettsia bellii* in cell culture from the tick *Amblyomma aureolatum* in Brazil. *Ann N Y Acad Sci* 2006;**1078**:523–9.
Pugsley AP. *Escherichia coli* K12 strains for use in the identification and characterization of colicins. *J Gen Microbiol* 1985;**131**:369–76.
Rachek LI, Tucker AM, Winkler HH *et al.* Transformation of *Rickettsia prowazekii* to rifampin resistance. *J Bacteriol* 1998;**180**:2118–24.
Radulovic S, Feng HM, Morovic M *et al.* Isolation of *Rickettsia akari* from a patient in a region where Mediterranean spotted fever is endemic. *Clin Infect Dis* 1996;**22**:216–20.

Ragonese H, Haisch D, Villareal E *et al.* The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at oriT through an interaction with TraI. *Mol Microbiol* 2007;**63**:1173–84.
Randolph SE. Transmission of tick-borne pathogens between co-feeding ticks: Milan Labuda's enduring paradigm. *Ticks Tick Borne Dis* 2011;**2**:179–82.
Rao Q, Wang S, Zhu DT *et al.* Draft genome sequence of *Rickettsia* sp. strain MEAM1, isolated from the whitefly *Bemisia tabaci*. *J Bacteriol* 2012;**194**:4741–2.
Raoult D, Paddock CD. *Rickettsia parkeri* infection and other spotted fevers in the United States. *N Engl J Med* 2005;**353**:626–7.
Reed SC, Lamason RL, Risca VI *et al. Rickettsia* actin-based motility occurs in distinct phases mediated by different actin nucleators. *Curr Biol* 2014;**24**:98–103.
Reif KE, Macaluso KR. Ecology of *Rickettsia felis*: a review. *J Med Entomol* 2009;**46**:723–36.
Renvoisé A, Joliot AY, Raoult D. *Rickettsia felis* infection in man, France. *Emerg Infect Dis* 2009;**15**:1126–7.
Saisongkorh W, Robert C, La Scola B *et al.* Evidence of transfer by conjugation of type IV secretion system genes between *Bartonella* species and *Rhizobium radiobacter* in amoeba. *PLoS One* 2010;**5**:e12666.
Sakurai M, Koga R, Tsuchida T *et al. Rickettsia* symbiont in the pea aphid *Acyrthosiphon pisum*: Novel cellular tropism, effect on host fitness, and interaction with the essential symbiont *Buchnera*. *Environ Microbiol* 2005;**71**:4069–75.
Sastre JI, Cabezon E, de la Cruz F. The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J Bacteriol* 1998;**180**:6039–42.
Schicht S, Schnieder T, Strube C. *Rickettsia* spp. and coinfections with other pathogenic microorganisms in hard ticks from northern Germany. *J Med Entomol* 2012;**49**:766–71.
Schriefer ME, Sacci JB, Dumler JS *et al.* Identification of a novel rickettsial infection in a patient diagnosed with murine typhus. *J Clin Microbiol* 1994a;**32**:949–54.
Schriefer ME, Sacci Jr JB, Taylor JP *et al.* Murine typhus: updated roles of multiple urban components and a second typhuslike rickettsia. *J Med Entomol* 1994b;**31**:681–5.
Sears KT, Ceraul SM, Gillespie JJ *et al.* Surface proteome analysis and characterization of surface cell antigen (Sca) or autotransporter family of *Rickettsia typhi*. *PLoS Pathog* 2012;**8**:e1002856.
Shapiro MR, Fritz CL, Tait K *et al. Rickettsia* 364D: a newly recognized cause of eschar-associated illness in California. *Clin Infect Dis* 2010;**50**:541–8.
Silverman DJ, Wisseman Jr CL, Waddell A. In vitro studies of *Rickettsia*-host cell interactions: ultrastructural study of *Rickettsia prowazekii*-infected chicken embryo fibroblasts. *Infect Immun* 1980;**29**:778–90.
Sivashanmugam A, Murray V, Cui C *et al.* Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Prot Sci* 2009;**18**:936–48.
Snyder JC, Wheeler CM. The experimental infection of the human body louse, *Pediculus*

humanus corporis, with murine and epidemic louse-borne typhus strains. *J Exp Med* 1945;**82**:1–20.
Soares HS, Barbieri AR, Martins TF *et al.* Ticks and rickettsial infection in the wildlife of two regions of the Brazilian Amazon. *Exp Appl Acarol* 2015;**65**:125–40.
Socolovschi C, Mediannikov O, Raoult D *et al.* The relationship between spotted fever group *Rickettsiae* and Ixodid ticks. *Vet Res* 2009;**40**:34.
Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc / .International Conf Intell Syst Mol Biol ; ISMBInternational Conf Intell Syst Mol Biol* 1998;**6**:175–82.
Sprong H, Wielinga PR, Fonville M *et al. Ixodes ricinus* ticks are reservoir hosts for *Rickettsia helvetica* and potentially carry flea-borne *Rickettsia* species. *Parasit Vectors* 2009;**2**:41.
Stothard DR, Clark JB, Fuerst P a. Ancestral divergence of *Rickettsia bellii* from the spotted fever and typhus groups of *Rickettsia* and antiquity of the genus *Rickettsia*. *Int J Syst Bacteriol* 1994;**44**:798–804.
Sunyakumthorn P, Bourchookarn A, Pornwiroon W *et al.* Characterization and growth of polymorphic *Rickettsia felis* in a tick cell line. *Appl Environ Microbiol* 2008;**74**:3151–8.
Taboada B, Ciria R, Martinez-Guerrero CE *et al.* ProOpDB: Prokaryotic Operon DataBase. *Nucleic Acids Res* 2012;**40**:D627–31.
Taboada B, Verde C, Merino E. High accuracy operon prediction method based on STRING database scores. *Nucleic Acids Res* 2010;**38**:e130.
Tamura A, Ohashi N, Urakami H *et al.* Classification of *Rickettsia tsutsugamushi* in a new genus, *Orientia* gen. nov., as *Orientia tsutsugamushi* comb. nov. *Int J Syst Bacteriol* 1995;**45**:589–91.
Tato I, Matilla I, Arechaga I *et al.* The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J Biol Chem* 2007;**282**:25569–76.
Tijsse-Klasen E, Jameson LJ, Fonville M *et al.* First detection of spotted fever group rickettsiae in *Ixodes ricinus* and *Dermacentor reticulatus* ticks in the UK. *Epidemiol Infect* 2011;**139**:524–9.
Toju H, Fukatsu T. Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: relevance of local climate and host plants. *Mol Ecol* 2011;**20**:853–68.
Toju H, Tanabe AS, Notsu Y *et al.* Diversification of endosymbiosis: replacements, co-speciation and promiscuity of bacteriocyte symbionts in weevils. *ISME J* 2013;**7**:1378–90.
Tomassone L, Nuñez P, Ceballos LA *et al.* Detection of "Candidatus *Rickettsia* sp. strain Argentina"and *Rickettsia bellii* in *Amblyomma* ticks (Acari: Ixodidae) from Northern Argentina. *Exp Appl Acarol* 2010;**52**:93–100.
Troughton DR, Levin ML. Life cycles of seven ixodid tick species (Acari: Ixodidae) under standardized laboratory conditions. *J Med Entomol* 2007;**44**:732–40.
Walker EM, Pittard J. Temperature-sensitive conjugation-defective F factor in *Escherichia coli*. *J Bacteriol* 1969;**100**:319–28.

Weinert LA, Welch JJ, Jiggins FM. Conjugation genes are common throughout the genus *Rickettsia* and are transmitted horizontally. *ProceedingsBiological Sci / R Soc* 2009;**276**:3619–27.
Weinert LA, Werren JH, Aebi A *et al.* Evolution and diversity of *Rickettsia* bacteria. *BMC Biol* 2009;**7**:6.
Wikswo ME, Hu R, Dasch GA *et al.* Detection and identification of spotted fever group rickettsiae in *Dermacentor* species from southern California. *J Med Entomol* 2008;**45**:509–16.
Willetts NS. Characterization of the F transfer cistron, *traL*. *Genet Res* 1973;**21**:205–13.
Williams SG, Sacci JB, Schriefer ME *et al.* Typhus and typhuslike rickettsiae associated with opossums and their fleas in Los Angeles County, California. *J Clin Microbiol* 1992;**30**:1758–62.
Wisseman Jr CL, Edlinger EA, Waddell AD *et al.* Infection cycle of *Rickettsia rickettsii* in chicken embryo and L-929 cells in culture. *Infect Immun* 1976;**14**:1052–64.
Wisseman Jr CL, Waddell AD. In vitro studies on rickettsia-host cell interactions: intracellular growth cycle of virulent and attenuated *Rickettsia prowazeki* in chicken embryo cells in slide chamber cultures. *Infect Immun* 1975;**11**:1391–404.
Woodard A, Wood DO. Analysis of convergent gene transcripts in the obligate intracellular bacterium *Rickettsia prowazekii*. *PLoS One* 2011;**6**:e16537.
Wozniak RA, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev* 2010;**8**:552–63.
Zavala-Castro JE, Zavala-Velázquez JE, Peniche-Lara GF *et al.* Human rickettsialpox, southeastern Mexico. *Emerg Infect Dis* 2009;**15**:1665–7.
Zavala-Velázquez JE, Ruiz-Sosa J a, Sánchez-Elias R a *et al. Rickettsia felis* rickettsiosis in Yucatán. *Lancet* 2000;**356**:1079–80.
Znazen A, Rolain JM, Hammami N *et al. Rickettsia felis* infection, Tunisia. *Emerg Infect Dis* 2006;**12**:138–40.