EVALUATION OF TARGETRON BASED MUTAGENESIS IN EHRlichia Chaffeensis

By

SHANZHONG GONG

BS., Southwest Agricultural University, China, 2004
MS., South China Agricultural University, China, 2007

A THESIS
Submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE
Department of Diagnostic Medicine/Pathobiology
College of Veterinary Medicine
KANSAS STATE UNIVERSITY
Manhattan, Kansas
2010

Approved by:

Major professor
Roman R. Ganta
ABSTRACT

*Ehrlichia chaffeensis* is an emerging tick-borne rickettsial pathogen that causes infection in people and several vertebrate animals. One of the striking features of *E. chaffeensis* infection is the prolonged persistence in its vertebrate and tick hosts. The mechanism of persistent infection and the reasons for the host immune system failure to clear the infection are not well understood. One hypothesis is that differential gene expression serves as an important adaptive mechanism used by *E. chaffeensis* in support of its continued survival in both tick and vertebrate hosts. One way to test this hypothesis is by performing mutational analysis. However, the methods for introducing mutations in this pathogen have not yet been documented and are challenging, possibly due to its obligate, intraphagosomesal growth requirement. Recently, a novel gene mutation method called ‘TargeTron Gene Knockout System’ that is based on the modified group II intron insertion strategy has been developed. This method appears to be effective in creating mutations in a wide range of gram positive and gram negative bacterial organisms. The group II intron can be programmed for insertion into virtually any desired DNA target with possibly high frequency and specificity. In this study, I focus on creating mutations in *E. chaffeensis* using the TargeTron gene knockout system. I prepared modified group II intron constructs retargeting for insertion into three *E. chaffeensis* genes: Ech_0126 (a transcriptionally silent gene), macrophage-specific expressed gene (p28-Omp 19, Ech_1143) and tick cell-specific expressed gene (p28-Omp 14, Ech_1136). In support of driving the expression of the modified group II introns in *E. chaffeensis*, the pathogen-
specific high-expressing gene promoter \((tuf)\) was inserted upstream to the transcription start site. In addition, a chloramphenicol acetyltransferase gene with \(E.\ chaffeensis\ rpsl\) promoter was introduced for use as a selection marker. The constructs were then evaluated by transforming into \(E.\ chaffeensis\). Transformants with mutations, introduced in two of the three genes (Ech_0126 and Ech_1143), were identified by PCR and Southern blot methods. Although the mutants are detectable for up to 48 hours, establishment of stable transformants remains to be challenging. The outcomes of this project will have important implications in defining the pathogenesis of \(E.\ chaffeensis\), particularly to assess the differences in the organism in tick and vertebrate hosts.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>CHAPTER ONE    LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER TWO    SCOPE OF THESIS</td>
<td>67</td>
</tr>
<tr>
<td>CHAPTER THREE   EVALUATION OF TARGETRON BASED MUTAGENESIS IN EHRlichia Chaffeensis</td>
<td>90</td>
</tr>
<tr>
<td>CHAPTER FOUR    MAJOR CONCLUSION</td>
<td>124</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Chapter I

Figure 1.1: Generalized mouthparts of ticks
Figure 1.2: One host, two hosts and three hosts of hard ticks
Figure 1.3: Multiple hosts of soft ticks
Figure 1.4: *Amblyomma americanum* (Lone star tick) and its distribution
Figure 1.5: White-tailed deer
Figure 1.6: Number of Ehrlichiosis cases (caused by *Ehrlichia chaffeensis*) reported to CDC by state health departments, 1999-2006
Figure 1.7: Age-specific incidence of Ehrlichiosis (caused by *Ehrlichia chaffeensis*), reported to CDC 2001-2002 through NETSS
Figure 1.8: Two morphological forms of *E. chaffeensis*
Figure 1.9: Generic secondary structure of group II intron RNA and domain structure of the intron-encoded proteins (IEP).
Figure 1.10: Mechanism of DNA target site recognition by the *Ll.LtrB* group II intron. (Jin Zhong *et al.*, 2003)

Chapter II

Figure 2.1: Overlapping PCR for group II intron retargeting

Chapter III

Figure 3.1: PCR for intron retargeting in gene 14, gene 19 and gene 126 of *E. chaffeensis*
Figure 3.2: Double digestion of the retargeted PCR products and pACD4K-C plasmids with Hind III and BsrG I.

Figure 3.3: Double digestion of constructed plasmids pACD4K-C-gene 14/gene 19/gene 126 with Hind III and BsrG I.

Figure 3.4: PCR for tuf promoter.

Figure 3.5: β-galactosidase activity of tuf promoter.

Figure 3.6: Hind III digestion of plasmids pACD4K-gene 14/gene 19/gene 126 and tuf promoter.

Figure 3.7: Confirmation of PACD4K-C-gene 14/gene 19/gene 126-tuf.

Figure 3.8: PCR for rpsL-CAT.

Figure 3.9: Double digestion with Mlu I.

Figure 3.10: Confirmation of pACD4K-C-gene 14/gene 19/gene 126-tuf-rpsL-CAT.

Figure 3.11: RNA gel of plasmids pACD4K-C-gene 14/gene 19/gene 126-tuf-rpsL-CAT.

Figure 3.12: RT-PCR, PCR for extracted RNA and PCR for plasmid DNA.

Figure 3.13: PCR strategies for detecting mutants.

Figure 3.14: Experiment 1, PCR assay for detecting insertion.

Figure 3.15: Experiment 2, PCR assay for detecting insertion.

Figure 3.16: Experiment 3, PCR assay for detecting insertion.

Figure 3.17: Experiment 1, southern blot for detecting insertion.

Figure 3.18: Southern blot for experiment 2.

Figure 3.19: Southern blot for experiment 3.
LIST OF TABLES

Chapter I

Table 1.1  Biologic vectors and vector-borne diseases

Table 1.2 Some of Rickettsialea Diseases

Chapter III

Table 3.1  Primers in this study
ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my major professor, Dr. Roman R. Ganta, who supported me throughout my thesis with his patience and knowledge. His invaluable assistance and guidance provides me with an excellent atmosphere for doing research. Under his supervision, our lab looks more likely a family. I attribute the level of my Masters degree to his encouragement and effort and without him this thesis, too, would not have been completed or written. One simply could not wish for a better or friendlier supervisor.

I would also like to show my sincere gratitude to all members of my advisory committee: Dr. M. M. Chengappa, Dr. Thu Annelise Nguyen, and Dr. Jianfa Bai for their kind assistance and extensive evaluation of my M.S program. They were very beneficial in providing me with ideas to better my overall project. I truly appreciate the time and effort they dedicated to helping me in the completion of my program.

I would especially like to thank Chuanmin Cheng for her guidance as well as friendship. Her capabilities in molecular biology helped me in the completion of my project. I would have never finished without her daily help and continuous support. She is more likely an elder sister than a lab manager, and I will always appreciate her help.
A special thanks also goes to my lab members: Huitao Liu, Bonto Faburay, Chakravarthy Katragadda, and Gangadaar Thotakura. To Dr. Huitao Liu, I will always be grateful for your great help and friendship. To Dr. Bonto, I would like to thank for your encouragement and support all the time. To Chakravarthy Katragadda, and Gangadaar Thotakura, thank you so much for creating an easygoing and cheerful laboratory atmosphere.

I would also like to thank Dr. Latitha Peddireddi, Dr. Geo-Moon Seo and Dr. Kamesh Reddy Sirigireddy for their encouragement. My sincere thanks also goes to my friends: Chuanmin, Huitao, Xianran, Haiyan, Yanru, Xiaofei, Qingyu and Pengzhao, for their moral support in everything.

This project is funded by the grant from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, USA (AI070908).

Shanzhong Gong
DEDICATION

This dissertation is dedicated to my parents, all my teachers, and to all of my friends who believed in me, and provided love and support with out which I would have been unable to complete this project.
Chapter I

LITERATURE REVIEW AND SIGNIFICANCE
Vector-Borne Diseases

Vector-borne diseases are the result of pathogen transferring from a vector to a vertebrate host, which caused by arthropod vectors including both insect (especially mosquitoes) and arachnid (ticks and mites). Mosquito-borne diseases include human malaria, filariasis, yellow fever, dengue fever and arthropod-borne viral encephalitis; tick-borne diseases include Rocky Mountain spotted fever, tularemia, Colorado tick fever, Q fever and relapsing fever; and mite-borne diseases include scrub typhus and scabies. In addition, crustacean, including copepods, crabs and prawns, also could transmit pathogens and cause diseases. For example, vibrio cholerae can persist in crabs for several weeks and multiply itself rapidly in free-living or refrigerated animals (Nalin et al., 1979). Vibrio cholerae can also colonize on the surface of copepods, and through them, it infects fish and human (Colwell et al, 1992; Estrada et al, 1996). Trematodes also harbor pathogens to serve as invertebrate hosts. For example, flukes host Neorickettsia species (Fukuda et al., 1981). In addition, Snails also serve as vector for transmitting disease agents. For example, they serve as intermediate hosts for Schistosomes and Angiostronglus Cantonensis and cause eosinophilic meningitis in human (Lv S et al., 2009).

Classification of the vector-borne diseases includes mechanical and biological vectors. Mechanical vectors are the vectors that simply carry pathogen from one host to another. For example, houseflies acquire bacteria on their appendages when they land on cow dung, and transfer the bacteria to humans when such flies land on food consumed by people. In this type of transmission, pathogen does not enter into the body of a
mechanical vector. Contrary to mechanical vectors, biological vectors usually harbor pathogens inside of their body and transmit to new hosts in a very active way, usually through biting. Pathogens develop and multiply themselves inside the body of biological vectors before becoming infective. Biological vectors play critical role in the life cycle of a pathogen. Many blood-feeding arthropods, such as mosquitoes, ticks, fleas and lice, are biological vectors. Some important biological vectors and vector-borne diseases are listed in Table 1.1.

Table 1.1 Biologic vectors and vector-borne diseases

<table>
<thead>
<tr>
<th>Vector</th>
<th>Scientific name/group</th>
<th>Disease</th>
<th>Disease pathogen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Insects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body louse</td>
<td><em>Pediculus humanus</em></td>
<td>Epidemic typhus</td>
<td><em>Rickettsia prowazekii</em></td>
</tr>
<tr>
<td></td>
<td><em>corporis</em></td>
<td>Epidemic relapsing fever</td>
<td><em>Borrelia recurrentis</em></td>
</tr>
<tr>
<td>Cone nose bugs</td>
<td><em>Triatoma</em> spp</td>
<td>Chagas’ Disease</td>
<td><em>Trypanosoma cruzi</em></td>
</tr>
<tr>
<td>Reduviid bugs</td>
<td><em>Rhodnius</em> spp</td>
<td>African Sleeping Sickness</td>
<td><em>Trypanosoma brucei</em></td>
</tr>
<tr>
<td>Assassin bugs</td>
<td><em>Panstrongylus</em> spp</td>
<td></td>
<td><em>rhodesiense</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>T. b. gambiense</em></td>
</tr>
<tr>
<td>Tsetse flies</td>
<td><em>Glossina</em> spp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horse flies</td>
<td><em>Chrysops</em> spp</td>
<td>Leishmaniasis</td>
<td><em>Loa loa</em></td>
</tr>
<tr>
<td>Sandflies</td>
<td><em>Phlebotomus</em> spp</td>
<td>Sandfly Fever</td>
<td><em>Leishmania</em> spp</td>
</tr>
<tr>
<td></td>
<td><em>Lutzomyia</em> spp</td>
<td>Oroya Fever</td>
<td>Arbovirus</td>
</tr>
<tr>
<td>Black flies</td>
<td><em>Simulium</em> spp</td>
<td>River blindness</td>
<td><em>Bartonella bacilliformis</em></td>
</tr>
<tr>
<td>Biting midges</td>
<td><em>Culicoides</em> spp</td>
<td>Filariasis</td>
<td><em>Mansonella</em> perstans</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Oropouche Fever</td>
<td><em>M. ozzardi</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Arbovirus</td>
</tr>
<tr>
<td>Mosquitoes</td>
<td>Anopheles spp</td>
<td>Malaria</td>
<td>Plasmodium spp</td>
</tr>
<tr>
<td>------------</td>
<td>---------------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>Lymphatic filariasis</td>
<td>Wuchereria bancrofti</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Onyog’nyong</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td>Culex spp</td>
<td>Filariasis</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oropouche</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ross River</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VEE.</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WEE</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>St Louis enceph</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jap. B. enceph</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td>Aedes spp</td>
<td>Yellow fever</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dengue</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rift Valley Fever</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ross River</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chikungunya</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EEE</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lymphatic filariasis</td>
<td>Brugia spp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wuchereria bancrofti</td>
<td></td>
</tr>
<tr>
<td>Psorophora spp</td>
<td>VEE.</td>
<td>Wuchereria bancrofti</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mayaro</td>
<td>Wuchereria bancrofti</td>
<td></td>
</tr>
<tr>
<td>Haemagogus spp</td>
<td>Yellow Fever</td>
<td>Wuchereria bancrofti</td>
<td></td>
</tr>
<tr>
<td>Mansonia spp</td>
<td>Lymphatic filariasis</td>
<td>Brugia spp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wuchereria bancrofti</td>
<td></td>
</tr>
<tr>
<td>Coquillettidia spp</td>
<td>EEE</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td>Ochlerotatus spp</td>
<td>Ross River</td>
<td>Arbovirus</td>
<td></td>
</tr>
<tr>
<td>Fleas</td>
<td>Xenopsylla spp</td>
<td>Bubonic plague</td>
<td>Yersinia pestis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Murine typhus</td>
<td>Rickettsia typhi</td>
</tr>
</tbody>
</table>

**Mites**

Rat mite | Leptotrombidium sp | Scrub typhus | Rickettsia tsutsugamushi

**Ticks**

Tampan | Ornithodoros spp | Endemic Relapsing fever | Borrelia duttonii
Despite the improvement of public health infrastructure in recent years, vector-borne diseases remain a very important issue for human and animal health. It is reported that more than half of total morbidity and mortality in people during the last four centuries resulted from vector-borne diseases (Gubler et al., 1998). The cost associated with vector-borne diseases is very huge because of lacking effective prevention and control. For example, malaria, a serious and potentially fatal disease caused by female mosquitoes of the genus *Anopheles*, causes 300-500 million clinical episodes and 1.2 million deaths each year (WHO, 2009). At the end of 2004, an estimated 3.2 billion people in 107 countries and territories (about 50% of the world population) live in areas at risk of malaria transmission (WHO, 2005). The majority cases of malaria occur in the world’s poorest countries, such as sub-Saharan African countries (Harrus et al., 2005). According to the World Health Report 2005, it also serves as one of the leading causes of death of
children in the developing countries (ranked 5th), and young children under the age of five or pregnant women are the most vulnerable group for this disease.

The economic impact of vector-borne diseases, which influence human health, is huge as well. For example, the economic impact of malaria in the developing countries accounts for 2-3% citizen income (Chima et al., 2003; Russel, 2004). The direct costs combined with other indirect costs cause greater impact. For example, the direct mean cost of malaria and other diseases in Nigeria amount to 7% of the household monthly income (Onwujekwe et al., 2000). However, if combined with indirect costs, they are responsible for more than 75% of the household income (Russel, 2004). Lyme disease is another example. Although it is not fatal, the mean annual incidence for this tick-borne disease in the USA during 1993–1996 is 4.73 cases per 100,000 populations. The expected national expenditure over 5 years for therapeutic interventions is about 2.5 billion dollars (Maes et al., 1998). Chagas disease (American trypanomiasis) is a gradually decreased disease in South and Central America. In Brazil, an estimated 516 million dollars are invested in the Chagas disease (kissing bug transmitted) control program between 1975 and 1995, of which 78% were spent on vector control. The control program is reported to be cost-effective with each one dollar spent on vector control resulting in 2.01 dollars of savings (Hoetz et al., 2004).

Vector-borne diseases also have enormous economic impact on agricultural animals. In the 1980s, it is estimated that hemoparasitic diseases of cattle could cost an annual global economic loss of 7 billion dollars (Brown, 1997). In the early 1990s, tick borne bovine babesiosis and anaplasmosis costed an annual economic loss of 875 -1365 million dollars
only in Latin America (James, 1992; Brown, 1997). In Africa, the vector borne diseases are also very severe. For example, tick transmitted heartwater disease is estimated to cost an economic loss of 5.6 million dollars each year in Zimbabwe (Mukhebi et al., 1999).

Introduction of tick and tick-borne diseases

Biology of Ticks

Tick is a common name of small arachnids that belong to the subphylum *Chelicerata*, class *Arachnida* and order *Ixodidea* (Sonenshine, 1991). It is an obligate blood sucking arthropods which have nearly 850 tick species (Oliver, 1989; Sonenshine, 1991). According to morphological differences, ticks could be classified as three families: *Ixodidae, Argasidae* and *Nuttalliellidae*. Ticks from family *Ixodidae* usually with tough sclerotized plate on their dorsal body surface, and thus they are called hard tick. Family *Argasidae* includes five genera and approximately 170 species were identified. Usually, they have a leathery cuticle on the dorsal body surface, and thus they are called soft ticks. Family *Nuttalliellidae* is confined to southern parts of Africa from Tanzania to Namibia and South Africa. This family includes only one species- *Nuttalliella namaqua* (Keirans et al., 1976), which is a morphologically intermediate between the families *Argasidae* and *Ixodidae*. In general, ticks belonging to family *Ixodidae* and family *Argasidae* serve as vectors for transmitting infectious agents to animals and people.

The mouthparts of ticks include three visible components (Figure 1.1): two outside jointed parts are highly mobile palps and between them are paired chelicerae. The chelicerae serve as protector for the center rod-shaped structure named hypostome. It is
the hypostome that has beak-like projection which could enter the skin of the host. Some backward directed projection prevents easy removal of attached tick from the host. In addition, they also secret a cement-like substance from salivary glands and this substance glues the body of the tick with the host. When feeding completed, this substance will dissolve automatically. The difference between family *Ixodidae* (hard ticks) and family *Aragasidae* (soft ticks) is that the mouthparts of hard ticks are visible from above, but the mouthparts of soft ticks are not readily visible if looked from the same position.

![Generalized mouthparts of ticks.](http://extension.entm.purdue.edu/publichealth/images/tick/tick02.gif)

The most important difference between hard ticks and soft ticks lies on their feeding styles and life cycles. Hard ticks usually are slow-feeders. They stay in the body of host for a few days, and take a large volume of blood meal each time, usually about 200-300 times of their unfed body weight. For soft ticks, they usually stay in the body of host for a few minutes or at most a few hours, and take a small volume of blood meal with only 5-10 times of their unfed body weight. Compared to hard ticks, soft ticks are fast feeder.
The life cycle of ticks includes four stages: egg, larva, nymph and adult. Eggs hatch into six legged larva stage, and then larva molt into nymphaal stage after taking a blood meals. Hard ticks and soft ticks are very different at this stage. Hard ticks only have one nymphaal stage. After nymphaal stage, they take another blood meal and develop into the adult stage. After feeding one more time, female hard ticks lay a batch of thousands of eggs and then die. For soft ticks, they molt into a multiple nymphaal stages and feed several times during each life stage. Unlike hard ticks, female soft ticks lay multiple small batches of eggs between blood meals and thus the time of completing their whole life cycle is much longer than that of hard ticks, sometimes lasting for more than several years. Also, soft ticks require multiple hosts for developing into their adult stages (Figure 1.3), but hard ticks have less than three hosts all their life (Figure 1.2).

Figure 1.2 One host, two hosts and three hosts of hard ticks.

(Source: http://entomology.ucdavis.edu/faculty/rbkimsey/tickbio.html)
Ticks as vectors

Ticks are known as vectors for transmitting a diverse group of infectious agents such as protozoa, bacteria and virus to vertebrate animals. For example, protozoan parasite *Babesia microti* transmitted by *Ixodes dammini* ticks caused infection of human babesiosis in the USA (Rodgers *et al.*, 2007). *Borrelia burgdorferi* transmitted by *Amblyomma americium* ticks caused Lyme diseases to people (CDC, 2001). *Francisella tularensis* transmitted by *Dermacentor* and *Haemophysalis* species ticks caused infection of tularemia in people (Gordon *et al.*, 1983). *Rickettsia rickettsia* transmitted by *Dermacentor* species ticks and cause Rocky Mountain spotted fever in the USA (Weber *et al.*, 1991). In addition, *Ehrlichia* (*E. chaffeensis* and *E. ewingii*) and *Anaplasma phagocytophilum* transmitted by *Amblyomma* and *Ixodes* ticks caused *Ehrlichiosis* in
people (Dumler et al., 1998; Alciati et al., 2001; Anderson et al., 1993; Bakken et al., 2000; Buller et al., 1999; Dumler et al., 1998). Tick-born encephalitis virus (TBEV) transmitted by *Ixodes* ticks caused encephalitis and hemorrhagic fever in people in many parts of the world (Mitzel et al., 2007).

**Tick-borne Rickettsiales Diseases**

Rickettsiales, transmitted by tick, lice, mites, fleas and flukes, include a group of obligate intracellular bacteria and most pathogen organisms causing fatal and chronic infections in people and vertebrate animals. Tick-borne Rickettsiales diseases are important human and animal diseases. According to the analysis of 16S rRNA, GroESL and several other surface protein gene sequences, they can be divided into two families: *Rickettsiaceae* and *Anaplasmataceae* (Dumler et al., 2001; Fournier et al., 2003).

The family *Rickettsiaceae* includes two genera: *Rickettsia* and *Orientia*. The genus *Rickettsia* has twenty-five species, whereas the genus *Orientia* has only one species- *O. tsutsugamushi* (Dumler et al., 2001). The genus *Rickettsia* can also be subdivided into two groups based on the genotypic and phenotypic similarities: the spotted fever group *rickettsiae* (SFGR) and typhus group *rickettsiae* (TGR) (Raoult et al., 1997; Dumler et al., 2001). SFGR is mainly transmitted by ticks and TGR is transmitted by flea and lice (Walker, 1989; Parola et al., 2001; Raoult et al., 1997; Uchiyama, 2005). Some important diseases of SFGR include *Rocky Mountain spotted fever* (Western hemisphere), *Rickettsialpox* (USA and former Soviet Union), *Boutonneuse fever* (Mediterranean countries, Africa, Southwest Asia and India), *Siberian tick typhus* (Siberia,
Mongolia, northern China), *Australian tick typhus* (Australia), *Oriental spotted fever* (Japan) and *African tick-bite fever* (South Africa). Some important diseases of TGR include *Epidemic, recrudescent and sporadic typhus* (worldwide), and *Murine* (endemic) *typhus* (worldwide). The species *O. tsutsugamushi* of the genus *Orientia* is mainly transmitted by mites (Rosenberg, 1997; Watt et al., 2003). A disease named scrub typhus formerly known as *R. tsutsugamushi* has also been reclassified into the genus.

The family *Anaplasmataceae* includes four genera: *Anaplasma, Ehrlichia, Neorickettsia* and *Wolbachia*. The genera *Anaplasma and Ehrlichia* are mainly transmitted by Ixodid ticks; the genus *Neorickettsia* is mainly transmitted by flukes (Dumler et al., 2001; Rikihisa, 2003; Hechmy et al., 2003; Raoult et al., 1997; Watt et al., 2003); and the genus *Wolbachia* is endosymbionts in arthropods and nematodes (Bandi et al., 1998; Dumler et al., 2001; Jeyaprakash et al., 2000; Werren et al., 2000). Some important tick-borne diseases of the genus *Anaplasma and Ehrlichia* in people include: human granulocytic *anaplasmosis, ewingii ehrlichiosis*, and *monocytic ehrlichiosis*, which are caused by *Anaplasma phagocytophilum, Ehrlichia ewingii*, and *Ehrlichia chaffeensis*, respectively. The genus *Neorickettsia* has three species reported pathogenic: *helminthoea, risticii* and *sennetsu* (Madigan et al., 2000; Rikihisa et al., 1991; Brown et al., 1972). The genus *Wolbachia* is non-pathogenic symbiotic bacteria resided in insects, spiders, mites, ticks and filarial nematodes (Stouthamer et al., 1999). Some of *Rickettsiales* diseases are listed in Table 1.2 and some of well-known tick-borne *Rickettsiales* are discussed below.
Table 1.2 Some of Rickettsiae\textsuperscript{a} Diseases

<table>
<thead>
<tr>
<th>Rickettsia species</th>
<th>Disease</th>
<th>Vectors</th>
<th>Hosts</th>
<th>Geographic distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Typhus group:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. prowazekii</td>
<td>Epidemic typhus</td>
<td>Human body</td>
<td>lice, Humans</td>
<td>Worldwide</td>
</tr>
<tr>
<td></td>
<td>Recrudescent typhus</td>
<td>None, Lice, fleas</td>
<td>Humans, Flying squirrels</td>
<td>Worldwide, Eastern USA</td>
</tr>
<tr>
<td>R. typhi</td>
<td>Murine typhus</td>
<td>Fleas</td>
<td>Rodents</td>
<td>Worldwide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fleas</td>
<td>Opossums</td>
<td>USA</td>
</tr>
<tr>
<td>R. felis</td>
<td>Murine typhus like</td>
<td>Fleas</td>
<td>Opossums</td>
<td>USA</td>
</tr>
<tr>
<td><strong>Spotted Fever group:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. rickettsii</td>
<td>Rocky Mountain</td>
<td>Ticks</td>
<td>Small mammals, dogs, rabbits, birds</td>
<td>North &amp; South America</td>
</tr>
<tr>
<td></td>
<td>spotted fever</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. conorii</td>
<td>Boutonneuse fever</td>
<td>Ticks</td>
<td>Rodents, dogs</td>
<td>Africa, Southern</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Europe, India</td>
</tr>
<tr>
<td>R. sibirica</td>
<td>North Asia tick typhus</td>
<td>Ticks</td>
<td>Rodents</td>
<td>Eurasia, Asia</td>
</tr>
<tr>
<td>R. japonica</td>
<td>Japanese spotted fever</td>
<td>Ticks</td>
<td>Rodents, dogs</td>
<td>Japan</td>
</tr>
<tr>
<td>R. australis</td>
<td>Queensland tick typhus</td>
<td>Ticks</td>
<td>Rodents</td>
<td>Australia</td>
</tr>
<tr>
<td>R. akari</td>
<td>Rickettsialpox</td>
<td>Mites</td>
<td>House mice, rats</td>
<td>Worldwide</td>
</tr>
<tr>
<td><strong>Ehrlichioses group:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. chaffeensis</td>
<td>Human monocytic ehrlichiosis</td>
<td>Ticks</td>
<td>Humans, deer</td>
<td>USA, Europe</td>
</tr>
</tbody>
</table>
Anaplasma phagocytophilum  Human granulocytic anaplasmosis  Ticks  Humans, deer rodents  USA, Europe

Ehrlichia ewingii  Human ewingii ehrlichiosis  Ticks  Humans, deer rodents  USA, Europe

Neorickettsia sennetsu  Sennetsu ehrlichiosis  Trematode  Humans  Japan, Malaysia

Others:

Orientia tsutsugamushi  Scrub typhus  Mites  Rodents  Asia, Indian subcontinent, Australia

*Not listed are R. helvetica, R. honei, and R. slovaca.*

( source: [http://www.cdc.gov/ncidod/eid/vol4no2/azad.htm#Table%201](http://www.cdc.gov/ncidod/eid/vol4no2/azad.htm#Table%201) )

**Rocky Mountain spotted fever**

*Rocky Mountain spotted fever* (RMSF) is the most severe tick-borne rickettsial disease in the USA, which caused by *Rickettsia rickettsii*. This rickettsial organism is primarily transmitted by American dog tick (*Dermacentor variabilis*) and Rocky Mountain wood tick (*Dermacentor andersoni*). In southern region of the United States, the main vectors are *Rhipicephalus sanguineus* (brown dog tick) and *Amblyomma cajennense* (CDC, 2010). Humans of all ages are susceptible to RMSF. The pathogen, which spread by blood and the lymphatic system, infects the whole body and causes severe damage to endothelial cells lining the blood vessels of vital organs (Masters *et al.*, 2003). Although this disease is first discovered and named after Rocky Mountain area, recently few cases are reported from that area. About half of the cases occur in the south-Atlantic region of
the United States: Delaware, Maryland, Washington D.C., Virginia, West Virginia, North Carolina, South Carolina, Georgia, and Florida. The highest incidence rates have been found in North Carolina and Oklahoma of the USA (CDC, 2010).

The early clinical signs of RMSF are often nonspecific which may include fever, nausea, vomiting, severe headache, muscle pain, lack of appetite. Usually, the characteristic red, spotted (petechial) rash is not visible until the sixth day or later after the onset of symptoms. Severe cases may also involve infection to the respiratory system, central nervous system, gastrointestinal system, or renal system. The mortality of this disease ranges from 5% to 10% (Chapman et al., 2006), and most of the mortalities are associated with delay in diagnosis (Masters et al., 2003).

**Mediterranean spotted fever**

*Mediterranean spotted fever* (MSF), also known as boutonneuse fever, is a tick-borne rickettsial disease caused by *Rickettsia conorii* and transmitted by the dog tick *Rhipicephalus sanguineus* (Azad et al., 1998). MSF is reported in the Mediterranean countries including southern Europe, northern Africa, and part of Asia (Roverey et al., 2008). Because of increased travel to endemic areas, this disease has also been transmitted to other regions of the world including central Europe and southern and central Africa (Roverey et al., 2008). After infection, the pathogen invades vascular endothelial cells, causing endothelial cells injury and tissue necrosis. Like the other rickettsial diseases, the symptom is non-specific: fever, severe headache, and maculopapular rash (Sousa et al., 2003; Roverey et al., 2008). If recognized at the initial
stages, MSF can be treated effectively with antibiotics such as chloramphenicol, tetracycline or doxycycline (Roverey et al., 2008).

**Japanese spotted fever**

*Japanese spotted fever*, also known as *Oriental spotted fever*, is another rickettsiae disease which transmitted by tick family *Ixodidae* and caused by *Rickettsia japonica*. This disease was first reported in 1984 in Japan and only confined to the region of southwestern Japan, but currently it also has been reported existing in Korea (Ma et al., 2005). Main symptom of the disease includes fever, headache and the appearance of an eschar and rash, and severe cases might involve the damage to the central nervous system (Araki et al., 2002).

**North Asian tick typhus**

*North Asian tick typhus*, also called *Siberian Tick Typhus*, is another tick-borne rickettsiae disease which caused by the bacteria *Rickettsia sibirica*. Symptoms include a generalized maculopapular rash involving palms and soles, fever, and lymph node enlargement. It occurs in the north, central and east parts of Asia, and because of traveling to endemic area, it was also reported in North America (Lewin et al., 2003). This disease is rarely fatal and easily cured by treating with chloramphenicol.

**Flinders Island spotted fever**

*Flinders Island spotted fever* was caused by *Rickettsia honei* and transmitted by *Ixodes holocyclus* tick (Stewart, 1991). It was first reported from Flinders Island near Australia in
1991 and isolated later in 1998 (Stewart, 1991; Stenos et al., 1998). The pathogen proliferates within the endothelial cells of small blood vessels and causes vasculitis. The usual symptom of the diseases includes high fever, headache, muscle aches (*myalgia*), slight cough, and arthralgia without joint swelling. Maculopapular rash might occur 5 or 6 days after tick biting. The disease is rarely fatal and can be cured by treatment of tetracyclines.

**Australian Spotted Fever**

*Australian Spotted Fever*, also known as *Queensland Tick Typhus*, is caused by *Rickettsia australis* and transmitted by *Ixodes holocyclus* tick (Graves et al., 1993). Like Flinders Island spotted fever, the infectious agent proliferates within the endothelial cells of small blood vessels and induces vasculitis. This disease was first reported in 1946 in north Queensland (Stewart et al., 1991). The symptom involves fever, headache and muscle aches (*myalgia*). Severe cases might also involve stiff neck, nausea, vomiting and mental confusion (Unsworth et al., 2007). This disease is also rarely fatal and more quickly cured with treatment of antibiotics (CDC, 2007).

**Ehrlichia ewingii**

*Ehrlichia ewingii*, originally known as the pathogen causing granulocytic ehrlichiosis in canines, was recognized in 1999 as the infectious agent for granulocytic ewingii ehrlichiosis in humans (Buller et al., 1999). The main natural vector of *E. ewingii*- the lone star tick- *Amblyomma americanum* and its recent infection in humans have been reported from Missouri, Oklahoma, and Tennessee (Buller et al., 1999; Paddock et al.,
The symptom of the infection includes fever, headache and myalgia, and usually reported in cases that recently be exposed to ticks. The diagnosis of this disease is mainly through isolation of the agent from tissue culture (Childs et al., 1999), and identification by PCR assay with two regions of gene sets serving as primers-16S rRNA genes and groE genes (Breitschwerdt et al., 1998; Buller et al., 1999; Sumner et al., 2000). The disease can be effectively cured with antibiotics like doxycycline.

**Human Granulocytic Anaplasmosis**

*Human granulocytic anaplasmosis* (HGA), previously known as human granulocytic ehrlichiosis, is caused by a bacteria called *Anaplasma phagocytophilum* and mainly transmitted by deer tick- *Ixodes scapularis*. This disease was first identified in 1990 in a Wisconsin patient who died from a febrile disease after two weeks of tick bite (Chen et al., 1994). Although it is mostly reported from Northeastern and upper Midwestern regions of the United States where the *Ixodes* ticks are highly prevalent (Demma et al., 2006), it also appeared in Pacific western states of the USA (Richter, 1996; Barlough, 1997) and western Europe (Stedink, 1997), where *Ixodes pacificus* (the black-legged tick) and *Ixodes ricinus* (the sheep tick) are prolific. The pathogen targets at white blood cells called neutrophils. Symptom of this disease includes high fever, severe headache, malaise, anorexia, nausea, cough and nonspecific rash. The duration time may only last a couple of days, but may also persist longer if untreated. This disease is also rarely fatal and can be treated effectively with antibiotics such as doxycycline.

**Human Monocytic Ehrlichiosis**
Human monocytic ehrlichiosis (HME), caused by *Ehrlichia chaffeensis* and transmitted by the lone star tick *Amblyomma americanum*, is one of the most important tick-borne Rickettsiale diseases in the USA. The pathogen is maintained in nature mainly in dual hosts, white-tailed deer (*Odocoileus virginianus*) and the lone star tick *Amblyomma americanum* (Dawson *et al*., 1994; Lockhart *et al*., 1995). HME cases were reported from the south-central, southeastern, and mid-Atlantic states of the USA, which corresponds to the geographical distribution of the lone star tick (CDC, 2009). Some cases may also be reported from the outside of the expected regions of the lone star tick, because many travelers serve as vectors to carry the infectious agent outside of the endemic areas (CDC, 2009). The symptoms of this disease include fever, headache, malaise, and muscle aches. Some clinical signs of HME cases are easily to be separated from other diseases like HGA. For example, rash is easier to be found in HME cases (Dumler *et al*., 2005). Some severe cases of the diseases may involve infection to center nervous system. Diagnosis of human monocytic ehrlichiosis is mainly based on symptom and clinical signs of the disease and confirmatory laboratory test. Some indicatives of human monocytic ehrlichiosis in confirmatory laboratory test include low white blood cell count, low platelet count and elevated liver enzymes (CDC, 2009). Treatment with a tetracycline antibiotics and especially doxycycline has been shown effective for this disease (Hamburg *et al*., 2008). Much more details about HME will be discussed in the later paragraphs.
Ehrlichia chaffeensis, an emerging human infectious agent

Discovery

Ehrlichia chaffeensis, the causative agent of human monocytic ehrliciosis, is obligatory intracellular tick-transmitted rickettsiales bacteria in the family Anaplasmataceae of class Alpha Proteobacteria. The disease that the bacteria caused, which is called human monocytic ehrliciosis, was first found in April 1986 by a medical intern in Arkansas scanning the peripheral blood smear of patients (Maeda et al., 1987). Because the description of sustained tick-bitten experience by the patients, it was suspected as Rocky Mountain spotted fever (Fishbein, 1990; Maeda et al., 1987). Then, clinicians and scientists found that this infectious agent actually belongs to the genus Ehrlichia (Maeda et al., 1987). It was isolated in cell culture and characterized by molecular techniques at Fort Chaffee, Arkansas (Dawson et al., 1991), and because of that, it was named as Ehrlichia chaffeensis.

Background

This obligatory intracellular pathogen principally transmitted by the lone star tick Amblyomma americanum (Figure 1.4), and the major vertebrate mammalian host for this organism is the white-tailed deer (Figure 1.5) (Dawson et al., 1994; Lockhart et al., 1997; Parola et al., 2005; Anderson et al., 1993). Other vertebrates including humans, dogs, goats and coyotes may also be susceptible for serving as accidental hosts (Dugan et al., 2000; Yabsley et al., 2004; Dugan et al., 2005; Yu et al., 2008). The life cycle includes propagation of the pathogen in its tick vector and vertebrate hosts. Larval or nymphal ticks get infected by the bacteria while they are feeding the infected vertebrate host and
then they transmit the pathogen to another healthy vertebrate host. Till now, no records have suggested that this pathogen can be transovarially transmitted (Long et al., 2003). HME cases are reported mostly from places where the lone star ticks and white-tailed deer are abundant (from New Jersey to Kansas and Southward) (Standaert et al., 1999; Bakken et al., 1996; Bakken et al., 1998; Dumler et al., 1998; Whitlock et al., 2000; Standaert et al., 2000; Olano et al., 2003; Paddock et al., 2003; McNabb et al., 2008). Recently, the incidence of human monocytic ehrlichiosis in the USA has increased gradually (Figure 1.6) (Paddock et al., 2003; McNabb et al., 2008), with an average of 600 human cases of HME each year (McNabb et al., 2008).

Figure 1.4 Amblyomma americanum (Lone star tick) and its distribution.

(source: http://www.cdc.gov/ncidd/dvrd/Ehrlichia/Natural_Hx/Natural_Hx.htm)
Figure 1.5 White-tailed deer.

(source: http://fw.dpnr.gov.vi/education/AndyWebPage/Pictures/Mammals/digiscope029.jpg)

Figure 1.6 Number of Ehrlichiosis cases (caused by *Ehrlichia chaffeensis*) reported to CDC by state health departments, 1999-2006.

(Source: CDC National Electronic Telecommunications System for Surveillance (NETSS) data)
Clinical signs

The clinical signs usually appear in 1-4 weeks after tick-bite (Parola et al., 2005). The symptom of this disease ranges from asymptomatic to mild flu-like symptoms, and sometimes might deteriorate into severe life threatening situation. The most common symptom includes malaise, fever, headache, muscle aches, chills, nausea and lymphadenopathy (Paddock et al., 2003; Dumler et al., 2007). Sometimes, it might also include cough, pharyngitis, swollen lymph nodes, vomiting, and possible changes in mental status. Unlike the Rocky Mountain spotted fever, only 30-40% of suspected cases reported having a rash associated with infection. Severe cases may include complications such as septic shock-like syndrome, meningitis, organ damage (Paddock et al., 2001; Stone et al., 2004; Dumler, 2005; Dumler et al., 2007). Leucopenia, thrombocytopenia, and an increase in hepatic transaminase levels are usually observable in the laboratory diagnosis (Paddock et al., 2003; Dumler et al., 2007). Although people of all ages are susceptible to the disease, it is reported that immune compromised people and elderly people have greater chance of getting infection (Figure 1.7) (Paddock et al., 2003; Gardner et al., 2003).
Diagnosis and treatment

Because other bacteria or virus also have similar symptom and clinical signs, definitive diagnosis during early stage of the disease is rather difficult. Thus, accurate description of tick-bite history and systematic laboratory diagnosis is very helpful to the preliminary confirmation of this disease (Dumler et al., 2007). Routine laboratory diagnostic methods for preliminary examination include detecting peripheral blood smear, performing IFA and PCR. Confirmatory examination usually involves cell culture of the pathogen (Anderson et al., 1991; Chapman et al., 2006; Dumler et al., 2007; Rikihisa et al., 2007). For prevention and clearance of the infection, treatment with tetracycline is proven to be very effective (Paddock et al., 2003). A derivative of tetracycline – doxycycline, which has been proved having better efficacy, is currently the drug of choice for treating HME (Dumler et al., 2007).
Molecular biology


16S rRNA

16S rRNA gene sequence has been reported from various *E. chaffeensis* isolates as an identical gene (Wen *et al.*, 1997), and because of it is highly conserved, the 16S rRNA is used for other species differentiation and phylogenetic classification (Anderson *et al.*, 1991; Dumler *et al.*, 2001). For example, PCR assays based on 16S rRNA gene and flowing PCR by RFLP is able to give a rapid discrimination between *A. marginale*, *A. centrale* and *A. ovis* (Vahid *et al.*, 2009). PCR assays based on 16s rRNA gene and electro spray ionization mass spectrometry can be used to rapidly detect and identify *Ehrlichia* species in blood (Mark *et al.*, 2010).
The 28 kDa outer membrane proteins (p28-Omp)

The 28 kDa outer membrane proteins of *E. chaffeensis* are encoded by a multigene locus containing 22 tandemly arranged paralogous genes (Reddy *et al.*, 1998; Ohasi *et al.*, 1998; Reddy *et al.*, 1999; Yu *et al.*, 2000; Cheng *et al.*, 2003; Singu *et al.*, 2005; Seo *et al.*, 2008). The molecular weights of mature proteins expressed from this multigene locus are predicted to be 28 kDa to 32 kDa, and because of that, it was named as the 28 kDa outer membrane proteins of *E. chaffeensis*. The amino acid sequence identity of these proteins is ranged from 20% to 83%. These arranged paralogous genes are separated by a fragment of 10 to 605 nucleotides non-coding intergenic sequences (Hotopp *et al.*, 2006). Three highly variable regions present in all 22 genes which are significantly different in each gene (Reddy *et al.*, 1998; Li *et al.*, 2001; Li *et al.*, 2002; Winslow *et al.*, 2000). They are hydrophilic in nature and contain the immunogenic B-cell epitopes (Reddy *et al.*, 1998; Winslow *et al.*, 2000; Li *et al.*, 2001; Li *et al.*, 2002).

The p28-Omp locus genes of *E. chaffeensis* isolates also contain considerable variations (Cheng *et al.*, 2003; Miura *et al.*, 2007). According to the sequence variation, *E. chaffeensis* can be classified into three groups (I, II, and III) (Cheng *et al.*, 2003). Ten isolates of *E. chaffeensis* recovered from human patients from various regions in the United States were used for this analysis. Group I contains three isolates: Arkansas, Osceola, and Lithonia. Group II contains five isolates: St. Vincent, Chattanooga, West Paces, Heartland, and Wakulla isolates. Group III contains two isolates: Liberty and Jax (Cheng *et al.*, 2003). In addition, the p28-Omp gene 18 and p28-Omp gene 15 were
found in group I isolates and group III isolates, respectively (Cheng et al., 2003). Variations existing in the genome of different *E. chaffeensis* isolates may have some connection for overcoming host immune responses for continued persistence in its dual hosts. This hypothesis remains to be tested.

*E. chaffeensis* p28-Omp multigene family in other species of *Ehrlichia* are also found (Gusa *et al.*, 2001; Ohashi *et al.*, 2001; van *et al.*, 2004; Crocquet, 2005; Zhang *et al.*, 2008). For example, in *E. ewingii* and *E. muris*, the p28-Omp multigene locus has 19, and 21 paralogous genes, respectively (van *et al.*, 2004; Crocquet, 2005; Zhang *et al.*, 2008). *E. canis* and *E. ruminantium* p28-multigene locus genus have different names, referred to as p30 Omp genes or MAP 1 genes (Ohashi *et al.*, 2001; van *et al.*, 2004). However, they are the paralogous of *E. chaffeensis* p28 Omp genes. The p-30 locus of *E. canis* includes 22 tandemly arranged paralogous genes (Ohashi *et al.*, 2001). Map 1/ p30 genes of *E. ruminantium* are 16 tandemly arranged paralogous genes (Brayton *et al.*, 1997; Ohashi *et al.*, 1998).

Another feature of the p28-Omp multigene family is the existence of genes that encode for a putative transcriptional regulator and a *secA* gene (Ohashi *et al.*, 2001; van *et al.*, 2004; Crocquet, 2006; Zhang *et al.*, 2008). The hypothetical transcriptional regulator of *E. ruminantium* is reported being polycistronically transcribed (van *et al.*, 2004). Recent studies from our laboratory also indicated the existence of transcriptional regulator in *E. chaffeensis* (Seo *et al.*, 2008; Sirigireddy *et al.*, 2005). The *secA* protein might be important in the process of translocation of bacterial outer membrane proteins to outer
membrane (Bernstein et al., 2000). The specific role of the putative transcriptional regulator and the secA gene in the p28-Omp multigene family of *E. chaffeensis* still needs to be studied.

It is reported that multiple genes of the p28-Omp multigene family are transcriptionally active in *E. chaffeensis* originating from the macrophages (Yu et al., 2000; Ohasi et al., 2001; Long et al., 2002; Unver et al., 2002; Cheng et al., 2003; Seo et al., 2008). However, the major expressed transcriptional protein in vertebrate cells is p28-Omp 19 (Cheng et al., 2003; Seo et al., 2008), and it is also reported that gene 14 of the p28-Omp multigene family is transcriptionally active in *E. chaffeensis* originating from tick cells or infected ticks (Unver et al., 2001; Unver et al., 2002). Similar differential expression for *E. canis* and *E. ruminantium* (Ohasi et al., 2001; Unver et al., 2001; Bekker et al., 2002; van et al., 2004).

Recent proteomic analysis from our laboratory further confirmed the predominant protein expression of p28-Omp gene 19 from vertebrate macrophages and the predominant protein expression of the p28-Omp gene 14 from tick cells (Singu et al., 2005; Singu et al., 2006; Seo et al., 2008). This host-specific predominant expression has also been reported in the p28-Omp locus homologues of *E. canis* (Unver et al., 2001; Felek et al., 2003; Singu et al., 2006). The vertebrate macrophages and tick cells host-specific predominantly expression might be an important mechanism to explain immune evasion and continued persistence of the pathogen in different hosts.
120 kDa surface protein

The 120 kDa surface protein is a potential adhesin of *E. chaffeensis* (Yu *et al.*, 2000). It is one of the major immunoreactive proteins for which antibodies were detected in most patient sera with HME infection and *E. canis* infected dogs (McBride *et al.*, 2000; Yu *et al.*, 1997; Yu *et al.*, 2000). The 120 kDa surface protein gene of Arkansas isolates of *E. chaffeensis* has four identical tandem repeats and each repeats is 240 bp in length. This gene feature shows great divergence between different *Ehrlichia* species, both in the number of the tandem repeats and the amino acid sequence of the repeats (Yu *et al.*, 1997). In addition, all of the major immunoreactive tandem repeat proteins are highly acidic due to a predominance of glutamate/aspirate and also they seems to be serine rich within tandem repeats of the proteins (Doyle *et al.*, 2006; Luo *et al.*, 2008). The 120 kDa surface proteins are differently expressed by dense-cored *E. chaffeensis* and are found on the surface of the organism (Popov *et al.*, 2000). It is reported that these surface proteins interact with important host cell targets and facilitate pathogen survival (Wakeel *et al.*, 2009). However, the specific role of this protein in pathobiology and protective immune response is still undefined.

Variable Length PCR Target gene (VLPT)

VLPT is another immunoreactive protein of *E. chaffeensis* and it can be used for detection and differentiation of *E. chaffeensis* strains (Sumner *et al.*, 1999). The VLPT of Arkansas strain of *E. chaffeensis* is 44 kDa in molecular mass and the open reading frame (ORF) contains four 90 bp imperfect tandem repeats (Sumner *et al.*, 1999). The number
of repeats is also not the same in different isolates of *E. chaffeensis* (Sumner et al., 1999, Cheng et al., 2003). For example, the Wakulla strain has a six-repeat version of the variable-length PCR target gene, whereas the Arkansas and Liberty strains have four repeats (Sumner et al., 1999). Compared to 120 kDa surface proteins, VLPT genes have more tandem repeats which make it more useful for strains differentiation (Yabsley et al., 2003). However, the exact role of VLPT in *E. chaffeensis* pathogenicity still remains to be known.

**Heat Shock Gene Operon (groESL)**

*E. chaffeensis* homologs of *Escherichia coli* heat shock gene operon (groESL) contains a stress-inducible promoter, followed by a GroES open reading frame which encodes a protein with molecular mass of 10-20 kDa, and a GroEL open reading frame which encodes a protein with molecular weight of 58-65 kDa (Sumner et al., 1993). GroES and GroEL genes are separated by noncoding sequences and these sequences vary in length in different bacteria (Sumner et al., 1993). GroEL has been called the bacterial common antigen and is related to the HSP 60 family of heat shock proteins (Rikihisa et al., 2007). Heat shock proteins of bacteria serve to protect cells under stressful condition and environment (Lindquist et al., 1988; Jenkins et al., 1988). However, the exact role of GroESL of *E. chaffeensis* is still not well studied.

**Immunologic aspects of *Ehrlichia chaffeensis* infection**

The possibility of getting infection by infectious agents depends on several factors, such as the age of an animal, or a specific host species, the virulence of a pathogen and the
immune status of the host. Usually, the pathogen intrusion might not indicate the necessary clinical signs of the disease, and the infection level may vary according to the host immune status. For example, the *E. chaffeensis* infection in humans is shown to be more severe in children, elderly and immune compromised people (Safdar et al., 2002; Paddock et al., 2003; Gardner et al., 2003), but very mild in the healthy adults. It is reported that *E. chaffeensis* can persist inside of healthy immune competent people, but it usually appears with no clinical sign (Davidson et al., 2001).

According the status of immune system of host, the responses for a pathogen intrusion may vary from a clinical disease, clearing of clinical signs and possibly even clearing the pathogen (Marshall et al., 2002; Yevich et al., 1995). For better understanding the host response of *E. chaffeensis*, several studies has been carried out since the discovery of the pathogen (Barnewall et al., 1994; Lee et al., 1996; Winslow et al., 1998; Winslow et al., 2000; Winslow et al., 2003; Ganta et al., 2002; Ganta et al., 2004; Ganta et al., 2007; Chapes et al., 2008). Recently, murine models has been proven very useful for illustrating the host immune mechanism against *E. chaffeensis* infection (Winslow et al., 1998; Winslow et al., 2000; Winslow et al., 2003; Ganta et al., 2002; Ganta et al., 2004; Ganta et al., 2007; Chapes et al., 2008).

In the early experiment of using murine models, immunocompetent mice clear the *E. chaffeensis* infection within 10-17 days post infection (Winslow et al., 1998), but immune deficient (SCID) mice who lack T- and B-cells fail to clear the infection of pathogen, and finally SCID mice exhibit severe clinical signs (Winslow et al., 1998). This proves that T-
cells and B-cells play vital role in clearing the infection of *E. chaffeensis* (Winslow *et al.*, 1998; Ganta *et al.*, 2002; Ganta *et al.*, 2004). In another experiment, major histocompatibility complex-II (MHC-II) deficient mice fail to clear the infection of *E. chaffeensis* (Ganta *et al.*, 2002; Ganta *et al.*, 2004), and the persistence of the pathogen in mice last for as long as 92 days (the longest time evaluated) (Ganta *et al.*, 2002; Ganta *et al.*, 2004). Similar results are obtained from toll-like receptor 4 (tlr4) deficient mice and it shows that the persistence of *E. chaffeensis* can last as long as 30 days (Ganta *et al.*, 2002; Ganta *et al.*, 2004). All the data indicate that major histocompatibility complex-II and toll-like receptor 4 are very important factor involving in the host immune response to *E. chaffeensis* infection.

Antibodies play a significant role in clearance of the pathogen. It is reported that SCID mice can be protected from a severe *E. chaffeensis* infection if injected with immune serum from *E. chaffeensis* infected immunocompetent mice (Winslow *et al.*, 2000). The protection time can last for up to 70 days if repeated injected with immune serum (Winslow *et al.*, 2000; Li *et al.*, 2001). Antisera against *E. chaffeensis* are predominantly made against outer membrane expressed proteins (Seo *et al.*, 2008). Although antibodies are known to play critical role in immune response to clear *E. chaffeensis* infection, they alone are not sufficient to clear the pathogen (Seo *et al.*, 2008). In particular, various subsets of T-cells and cytokines are needed for effective clearance of the infection (Li *et al.*, 2001; Winslow *et al.*, 2000).
Some subsets of T-cells have been proven to be important for immune response to *E. chaffeensis* infection (Ganta *et al.*, 2002; Ganta *et al.*, 2004; Bitsaktsis *et al.*, 2004; Bitsaktsis *et al.*, 2006; Nandi *et al.*, 2007; Ismail *et al.*, 2007; Bitsaktsis *et al.*, 2007; Chapes *et al.*, 2008). Helper T-cells deficient MHCII knockout mice can not clear the infection of *E. chaffeensis* for several months (Ganta *et al.*, 2002; Ganta *et al.*, 2004). The time of clearing the pathogen has been delayed for about two weeks in CD4+ helper T-cell deficient mice (Ganta *et al.*, 2002; Ganta *et al.*, 2004). This result indicates that helper T-cells are very important factor for host immune response for *E. chaffeensis*. It is also reported that γδ T-cells are found to be present in peripheral blood of *E. chaffeensis* infected human patients (Caldwell *et al.*, 1996; Caldwell *et al.*, 1996). The result indicates that γδ T-cells may be an important factor for immune response for this pathogen.

Other factors of host immune system might play minimum role for *E. chaffeensis* infection. For example, both CD4+ T-cells deficient mice and CD4+ T-cells deficient MHCII knockout mice exhibit a minimal CD8+ T-cells activity after *E. chaffeensis* infection (Ganta *et al.*, 2004); and β2 microglobulin (β2M, a structural component of MHCI)-knockout mice also exhibit a complete clearance of the infection of *E. chaffeensis* (Chapes *et al.*, 2008). The findings indicate that CD8+ T-cells are a minor factor in host immune clearance for the pathogen. This applies for NKT cells as well. The SCID mice that are deficient for T-cells and B-cells but contain functionally normal NKT cells exhibit persistent infection of *E. chaffeensis* (Winslow *et al.*, 1998). The NKT cells are
functionally normal, but *E. chaffeensis* is persistently existed in mice (Chapes *et al.*, 1993; Ganta *et al.*, 2002; Ganta *et al.*, 2004).

Some immunomodulating agents are involved in the clearance of the infection of *E. chaffeensis*. It is reported that suppression of several proinflammatory cytokine expression, including IL-1α, IL-4, IL-6, IL-12, IL-15, and IL-18 occurs in *E. chaffeensis* infected human monocytic cell lines (THP1 cells) (Zhang *et al.*, 2004). These cytokines are involved in early inflammatory response for the intrusion of the pathogen (Dinarello, 2000). Suppression of these cytokines will facilitate *E. chaffeensis* to escape from being killed by NKT cells and cytotoxic T-cells, and colonize inside the macrophages or monocytes (Dinarello, 2000). Cytokines IL-15 and IL-18 usually serve as activators for CD4⁺ T-cells to produce IFN-γ and activate macrophages mediated killing of the pathogen (Barnewall *et al.*, 1994; Liew *et al.*, 2002; Bitsaktsis *et al.*, 2004). Thus, cytokines are in part a critical factor in immune response for *E. chaffeensis* infection.

The entire immune system works in a network way, and if one subset of T-cells is absent, the other subsets of T-cells will compliment its function. The experiments of intracellular bacteria of *Mycobacterium tuberculosis* and *Francisella tularensis* illustrate this phenomenon. When CD4⁺ T-cells deficient mice are attacked by pathogens, other T-cells such as CD4⁻ T-cells, CD8⁻ T-cells, and natural killer T-cells will compliment for the CD4⁺ T-cell deficiency and work together to help clear the pathogens (Rahemtulla *et al.*, 1994; Cowley *et al.*, 2003; Cowley *et al.*, 2005; Mattner *et al.*, 2005). It is also reported that components of immune system including antibodies, macrophage activation,
involvement of several T-cell subtypes and cytokines work together to clear *E. chaffeensis* infection (Barnewall *et al.*, 1994; Lee *et al.*, 1996; Winslow *et al.*, 1998; Winslow *et al.*, 2000; Ganta *et al.*, 2002; Ganta *et al.*, 2004). Thus, network of different components of immune system serves as the effective solution for the immune system of the host in the clearance of *E. chaffeensis*.

**Electromicroscopy**

Under electromicroscopy, two morphological forms of *E. chaffeensis* can be observed residing within the intracellular parasitophorous vacuoles (morulae) in macrophages: dense-cored cells (DC) (0.4–0.6 μm) and reticulate cells (RC) (0.4–0.6 μm by 0.7–1.9) (Popov *et al.*, 1995; Zhang *et al.*, 2007) (Figure 1.8). DCs are smaller than RCs and they contain highly dense nucleotide inside; RCs contain uniformly dispersed nucleotide filaments and ribosome (Popov *et al.*, 1995). Some RCs form long projections of the cell wall and these projections protrude the cytoplasmic membrane into the periplasmic space, budding protoplast fragments into the periplasmic space (Popov *et al.*, 1995; Zhang *et al.*, 2007). Both DC and RC forms replicate themselves by binary fission inside the cells. In the beginning of their infectious cycle, usually 24 h after infection, DC usually transform into RC form. In the end of each cycle, the RCs transform back into DCs in the phagosome and they are released from monocytes or macrophages (Zhang *et al.*, 2007). This type of cycle, transforming from dense core to reticulate forms and then back to dense core forms, is very typical for the genera *Ehrlichia* and *Anaplasma* (Ashley *et al.*, 1975; Hodinka *et al.*, 1988; Wolf *et al.*, 2000; Wyrick, 2003). DCs and RCs are also
found in the phagosome of infected tick cells (Ganta et al., 2009). In tick cells, however, both DCs and RCs can exist within the same phagosome (Ganta et al., 2009).

![Figure 1.8](image)

**Figure 1.8** Two morphological forms of *E. chaffeensis* (Ganta et al., 2009).

**Pathogen persistence**

The pathogens of tick-transmitted diseases usually persist in their host for long time and cause severe infection. For example, *Borrelia burgdorferi* causes Lyme diseases in human, dogs and many other animals (CDC, 2010); protozoan parasite *Babesia microti* causes infection of human babesiosis in the USA (CDC, 2010); *Francisella tularensis* causes infection of tularemia in human (CDC, 2010); and *Rickettsia rickettsia* causes Rocky Mountain spotted fever to human (CDC, 2010). For an intracellular bacterium, like *E. chaffeensis*, the critical steps to establish effective infection usually include: attachment, internalization and intracellular multiplication (Rudel et al., 1992; Giron et al., 2002;
Okamoto et al., 2004, Zhang et al., 2004). *E. chaffeensis* lacks the specialized structures to attach and internalize, but it is reported that the 120 kDa surface protein appears to serve as adhesion (Popov et al., 2000). Nonetheless, it is unclear how this pathogen is phagocytosed by the host monocytes or macrophages, and it is also not clear how *E. chaffeensis* can overcome the phagosomal clearance and adapt itself to the host for keeping persistent infection. The mechanism of persistent infection and the reasons for the host immune system failure to clear the infection are not well understood.

One important hypothesis is that differential gene expression may be an important adaptative mechanism used by *E. chaffeensis* in support of its continued survival. Recent studies from our laboratory showed that the protein expression of *E. chaffeensis* from tick and vertebrate host cell environment has significant differences, especially the outer membrane proteins of the p28-Omp multigene family. Proteins of the gene 14 of the p28-Omp locus were found predominantly expressed in tick host, and proteins of gene 19 of the p28-Omp locus were found predominantly expressed in vertebrate host. Much of the fundamental knowledge of the pathogen’s persistence and its host specificity can be better understood by performing mutational analysis on the critical genes of the p28-Omp multigene family of *E. chaffeensis*.

Several methods have been documented and proven to be useful in creating effective gene mutations in both gram positive and gram negative bacteria, including in Rickettsial agents. For example, transposon-based mutagenesis has been applied for *Rickettsia* and *Anaplasma* species (Qin et al., 2004; Liu et al., 2007). Homologous recombination based
mutagenesis is extensively used in *Escherichia coli* and many other bacteria (Kowalczykowski *et al.*, 1994; Messerle *et al.*, 1997; Nowosielska *et al.*, 2006). However, the transformation in *E. chaffeensis* is complicated as it is an intracellular obligate pathogen and the host cell free bacteria have shorter life span (Paddock *et al.*, 2003). Methods like transposon based mutagenesis is usually targeted randomly, thus it is not appropriate to use it as the mutagenesis to knockout specific genes such as the genes of p28-Omp locus of *E. chaffeensis*. Although the homologous recombination based gene knockout system is a method specifically for targeting in one or several gene sites, this gene mutation method is not studied in detail in *E. chaffeensis*.

**Creating gene mutation in *E. chaffeensis***

Targetron gene knockout system (mobile group II introns) is a novel gene mutagenesis, which can be programmed to insert into any desired DNA target site with high specificity (Guo *et al.*, 2000; Karberg *et al.*, 2001). Mobile group II introns use a remarkable mobility mechanism mediated by a ribonucleoprotein (RNP) complex which contains the intron-encoded protein (IEP) and the excised intron RNA (Lambowitz *et al.*, 1999). Because most of the target specificity comes from base-pairing of the intron RNA to the target DNA sequence, it is possible to reprogram group II introns to insert into desired sites simply by modifying the intron RNA (Guo *et al.*, 1997). This feature combined with their very high insertion frequencies and specificity made it easy to create specific gene mutation on any organism. This method has been utilized successfully for *E.coli, Shigella flexneri, Salmonella typhimurium, Lactococcus lactis, Clostridium perfringens, Staphylococcus aureaus, and Francisella tularensis* (Yao *et al.*, 2006; Cui *et al.*, 2007;
John et al., 2007; Rodriguez et al., 2008). We believe that the Targetron gene knockout system may also be applied to study key genes of E. chaffeensis including p28-Omp multilocus genes.

**Introduction: Group II intron and Targetrons**

Group II intron is a class of intron existing in rRNA, tRNA, mRNA of organelles in fungi, plants, protists, and mRNA in bacteria (Dai et al., 2003). They were first discovered in the organelar mitochondrial and chloroplast genomes of lower eukaryotes and higher plants (Michel et al., 1989). The bacterial group II introns were identified about 17 years ago by PCR screens (Ferat et al., 1993). Then, genome sequencing revealed that group II introns are very common in both gram-negative and gram-positive bacteria (Dai et al., 2003). Later, group II introns have been found in Methanosarcinales (Toro et al., 2003; Dai et al., 2003; Rest et al., 2003). Recently, mobile group II introns have been used to develop a new type of gene-targeting method- Targetron Gene Knockout System (Lambowitz et al., 2004).

**Structures of Group II Introns**

Group II introns generally contain 6 domains (Domain I to Domain VI) (Figure 1.9), in which Domain I, Domain V, and Domain VI are involved the splicing reaction, Domain II and Domain III are involved in stabilization of proper folded structures, and Doman IV is involved in folding or catalytic activity. Domain I is the largest and most complex domain, in which the sequences of exon-binding sites 1 and 2 (EBS1 and EBS2) bind specifically to the sequences of intron-binding sites 1 and 2 (IBS1 and IBS2). Domain V is a small hairpin with a bulge in the stem and it contains a loop which is essential for
interaction with Domain I. It is suggested that Domain V is an essential catalytic component of the active ribozyme, and this domain can induce splicing reaction in special conditions (Qin et al., 1998). Domain VI contains a bulging “A” residue which is a branch point to the 5’ exon when splicing. Domain IV contains ORF expressing the intron-encoded protein (IEP) gene, and IEP is known to play an essential role in helping in vivo folding of the intron RNA into a catalytic active structure (Belfort et al., 1995; Zimmerly et al., 1995). The IEP includes four conserved domains: a N-terminal RT domain, a domain termed ‘X’ which is a putative RNA-binding domain associated with RNA splicing activity, a C-terminal DNA-binding domain, and a DNA-endonuclease domain. It is believed that the IEP is translated from a ‘pre-mRNA’ and later processed to the mature protein (Anziano et al. 1982; Carignani et al. 1983).

- 40 -
Retrohoming Mechanism

The mobility of group II introns is first demonstrated for the *S. cerevisiae* coxl-I1 introns (Meunier et al., 1990) and the related *Kluyveromyces lactis* coxl-I1 introns (Skelly et al., 1991). Then, it was reported that an efficient *E.coli* expression system for *Lactococcus lactis* Ll.LtrB intron has been constructed and group II introns could function as mobile genetic elements by inserting directly into target sites in double stranded DNA (Matsuura et al., 1997). This mobility is mediated by the multifunctional IEP, which shows reverse transcriptases (RTs), maturase, and DNA endonuclease activities (En) activities.

Retrohoming occurs through a target DNA-primed reverse transcription mechanism (Figure 1.10) which involves several sequential reactions. First, RNP initiates mobility by recognizing a specific double-stranded DNA target site, with both the IEP and base pairing of the intron RNA contributing to DNA target site recognition. Then, the IEP interacts with a few bases in the distal 5’ exon region via major groove and phosphate-backbone interaction, and triggers local DNA unwinding in the same time to facilitate the intron RNA base pairing with the adjacent 14–16 nt local DNA sequence. These base pairing interactions involve three short sequence elements in the DNA target sites: IBS2 and IBS1 and δ’, and they are recognized by three complementary sequence elements: EBS1 and EBS2 and δ in the intron RNA. After binding, the RNA component of the RNP cleaves the sense-strand precisely at the exon junction in the double-stranded DNA, and IEP uses its endonuclease domain (En) to cleave the antisense-strand at position +9 relative to the insertion site. The 3’ end of the cleaved antisense-strand is then targeted as
a primer for reverse transcription, generating a cDNA strand that is subsequently incorporated into the recipient DNA primarily by recombination mechanisms (Lambowitz et al., 2004). This mechanism can be used to develop the novel gene mutagenesis- Targetron Gene Knockout System.

Figure 1.10 Mechanism of DNA target site recognition by the Ll.LtrB group II intron (Zhong et al., 2003).

**Applications of Targetron Gene Knockout System**

Targetron Gene Knockout System is the commercial name of the technology, and it could be used to insert into any desired target DNA and knockout a number of genes. Due to its high efficiency and high specificity, Targetron Gene Knockout System has been successfully applied into gene mutation or modification in plentiful of bacterial species. Also, this technology significantly reduced the time needed for disrupting a gene, compared with conventional homologous recombination methods. For example, *Staphylococcus aureus*, which acquired antibiotic resistance, is a serious and increasing
clinical disease (CDC, 2001; CDC, 2005). In *Staphylococcus aureus*, hsa gene encodes a member of the HU family of bacterial histone-like proteins and seb encodes eterotoxin B (Yao et al., 2006). Creating site-specific mutations in this gram-positive bacterium targeting at the hsa and seb genes by using conventional homologous recombination is rather difficult. However, the hsa and seb genes were disrupted efficiently with Targetron Gene Knockout System (Yao et al., 2006). In higher organisms, group II introns were designed and selected to insert into HIV1 provirus and the human gene encoding CCR5, an important target site in anti-HIV therapy (Guo et al., 1997). The retargeted intron RNPs retained activity in human cells, inserting into plasmid-borne HIV1 and CCR5 target sites after liposome-mediated transfection (Guo et al., 1997).
Reference:


Dai L, Zimmerly S. 2003. ORF-less and reverse-transcriptase-encoding group II introns in archaebacteria, with a pattern of homing into related group II introns ORFs. RNA 9:14-19


Gross L. 1995. How the plague bacillus and its transmission through fleas were discovered: reminiscences from my years at the Pasteur Institute in Paris. PNAS. 92:7609–7611.


Lee, E. H. and Y. Rikihisa. 1996. Absence of tumor necrosis factor alpha, interleukin-6 (IL-6), and granulocyte-macrophage colony-stimulating factor expression but presence of IL-1beta, IL-8, and IL-10 expression in human monocytes exposed to viable or killed Ehrlichia chaffeensis. Infect. Immun. 64:4211-4219.


Chapter II

Scope of Thesis
Ehrlichia chaffeensis, an intracellular tick-transmitted rickettsial, is the agent of an infectious disease termed human monocytic ehrlichiosis (HME). The infection results in symptoms that vary from asymptomatic infection to severe life threatening disease with a mortality of about 3% (Paddock et al., 2003). Recently, the incidence of HME in the USA has increased (Paddock et al., 2003; McNabb et al., 2008) with an average of 600 reported cases each year (McNabb et al., 2008). The estimated cases may be significantly higher by at least four times more than the reported cases (CDC, 2007). One of the striking features of this disease causing agent is its dual host life cycle and prolonged persistence in both tick and vertebrate hosts. The mechanism of this persistent infection and the reasons for the host immune system failure to clear the infection is not clear.

There may be several ways by which the pathogen could persist. One important hypothesis is that differential gene expression may be an important adaptative mechanism used by E. chaffeensis in support of its continued survival in its dual hosts. Recent studies from our laboratory also support this hypothesis. Protematic analysis showed that the proteins of E. chaffeensis from tick and vertebrate host cell environment differed significantly. The differentially expressed proteins include many outer member proteins. An outer membrane protein from the p28-Omp mutigene locus gene p28-Omp 14 was found to be predominantly expressed in tick host and protein of the p28-Omp 19 gene was found to be the major expressed protein in vertebrate host environment. The differential expression was also confirmed in the bacteria under in vivo conditions in ticks and vertebrate hosts. Much of the fundamental knowledge of the pathogen’s persistence
and its host specificity may be better understood by performing mutational analysis on the critical genes of *E. chaffeensis*.

Studies to date on *E. chaffeensis* lead to considerable progress in identifying the differential expression of many proteins in a host cell specific manner. The recent studies also aided in understanding the host immune response needed for clearing the pathogen in a vertebrate host. In particular, studies suggest that *E. chaffeensis* can be cured rapidly if it originates from vertebrate macrophages, but the clearance may be delayed if the organism replicates in tick cell environment. These observations, together with host cell-specific differential expression, suggest that the altered pathogen gene expression in tick cell environment may provide an opportunity to persist in a vertebrate host. Similarly, the pathogen originating from vertebrate macrophages may persist in ticks. This may be important strategy for the pathogen’s continued life cycle in vertebrate and tick hosts. The above stated hypothesis is exciting, however, at this time we do not have molecular tools to test it. Importantly, there are no established methods to create mutations in *E. chaffeensis*, assessing the functional significance of differentially expressed protein.

Several methods have been documented and proven to be useful in creating effective gene mutations in both gram positive and gram negative bacteria, including in Rickettsial agents. For example, transposon based mutagenesis has been applied for *Rickettsia* and *Anaplasma* species (Qin *et al.*, 2004; Liu *et al.*, 2007). Homologous recombination based mutagenesis is extensively used in *Escherichia coli* and many other bacteria (Kowalczykowski *et al.*, 1994; Messerle *et al.*, 1997; Nowosielska *et al.*, 2006).
Nonetheless, these approaches have not been tested for *E. chaffeensis*. One of the challenging with *E. chaffeensis* is its obligate parasitic life and the pathogen has a limited survival outside a host cell. The focus of this study is to fill in this gap of knowledge by evaluating mutational methods. We opted to evaluate a relatively recently established method, ‘TargeTron Gene Knockout System’ for creating mutations in *E. chaffeensis*. This method utilizes the target specific insertions using modified group II intron. This method was effective in creating mutations in a wide range of gram positive and gram negative bacterial organisms. The justification for using this method is that the Group II intron can be programmed for insertion into virtually any desired DNA target with possibly high frequency and specificity. We focused this study with the objective to create mutations in *E. chaffeensis* at three genomic regions.
Materials and methods

1 Target site selection and primer ordering

The sequences of Ech_1136 (p28-Omp gene 14), Ech_1143 (p28-Omp gene 19) and Ech_0126 (p28-Omp gene 126) of *E. chaffeensis* downloaded from GenBank were used in preparing primers for retargeting group II intron with the aid of the TargeTron Design Site software (www.sigmaaldrich.com/targetron). The primer sets (IBS, EBS2, and EBS1d) that are likely to be most effective in the insertion mutations were ordered from Sigma Aldrich Company.

2 PCR for Group II intron re-targeting

2.1 Primer preparation

The IBS and EBS1d primers were diluted to 100 μM, and the EBS2 primer was diluted to 20 μM, in nuclease free water. These three primers (one set each for each gene retargeting) were mixed with the fourth universal primer, EBS universal primer, to prepare a mix for overlapping PCR as show in Figure 2.1:

```
Primer Mix

| 2 μl | IBS primer (100 μM) |
| 2 μl | EBS1d primer (100 μM) |
| 2 μl | EBS2 primer (20 μM) |
| 2 μl | EBS universal primer (20 μM) |
| 12 μl | Nuclease free water |
| 20 μl | Total volume |
```
2.2 PCR reaction set up and thermocycling

Amplification reactions were performed using the four primer mixes specific for each gene target and group II intron PCR template (Sigma, St. Louis, MO) by preparing the mixes and temperature cycles outlined below.

**PCR reaction mixture**

<table>
<thead>
<tr>
<th>Volume (μl)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>water (molecular biology reagent)</td>
</tr>
<tr>
<td>1</td>
<td>Four-primer mix (Lac Z control or specific primers for your target site)</td>
</tr>
<tr>
<td>1</td>
<td>Intron PCR template</td>
</tr>
<tr>
<td>25</td>
<td>JumpStart REDTaq Readymix</td>
</tr>
<tr>
<td>50</td>
<td>Total volume</td>
</tr>
</tbody>
</table>
Temperature cycles

<table>
<thead>
<tr>
<th>Step</th>
<th>Temp/Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>94 °C for 30 sec</td>
</tr>
<tr>
<td>30 cycles</td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C for 15 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>55°C for 30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C, 30 sec</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C, 2 min</td>
</tr>
<tr>
<td>Soak</td>
<td>4°C</td>
</tr>
</tbody>
</table>

2.3 Agarose gel electrophoresis

The PCR products were resolved in a 0.9 % agarose gel in 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA; final pH 8.0) containing 0.1 μg/ml of ethidium bromide by subjecting to 120 V of electricity for 1.5 h. The presence of 0.35 kb amplicons were identified with the aid of resolved molecular reagent markers. Agarose gel preparation and electrophoresis were performed by following instruction outlined in the volume 1 of the third edition of Molecular Cloning (Joseph et al., 2001).

2.4 Phenol purification of the PCR products and restriction enzyme digestions

PCR products were purified by following phenol:chloroform:isoamylalcohol and ethanol precipitation method (Joseph et al., 2001). Briefly, to 40 μl of PCR products, sodium acetate was added to a final concentration of 0.3 M. Two hundred μl of phenol (pH 8.0) was mixed to the PCR product and salt mix, vortexed and spun at 15,000 g for 5 min and aqueous layer was transferred to a clean 1.5 ml micro centrifuge tube. The organic extraction step was repeated two more times: one with phenol:chloroform:isoamylalcohol (25:24:1) and one with chloroform:isoamylalcohol (24:1). To the final resolved aqueous
layer, 500 μl of 100% ethanol was added and vortexed to mix the contents, and incubated at -20°C for 15 min followed by centrifugation at 15,000 g for 15 min. The supernatant was discarded and the DNA pellet was rinsed with 0.5 ml of 70% ethanol. The final pellet was air dried, resuspended in 20 μl TE buffer (pH 8.0) and stored at -20°C until use. Purified PCR products were digested with *Hind* III and *Bsr* G I (Note: These restriction sites were created by inserting the R.E. recognition sequence in the 5’ end of IBS, and 3’ end of EBS1d primers, respectively) under the conditions outlined in the table below.

<table>
<thead>
<tr>
<th>Volume</th>
<th>Substance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 μl</td>
<td>10X N.E. Buffer 2</td>
</tr>
<tr>
<td>8 μl</td>
<td>purified PCR products (about 200 ng)</td>
</tr>
<tr>
<td>1 μl</td>
<td><em>Bsr</em> G I</td>
</tr>
<tr>
<td>1 μl</td>
<td><em>Hind</em> III</td>
</tr>
<tr>
<td>0.2 μl</td>
<td>BSA (100 μg/ml)</td>
</tr>
<tr>
<td>8 μl</td>
<td>Nuclease free water</td>
</tr>
<tr>
<td>20 μl</td>
<td>Final volume</td>
</tr>
</tbody>
</table>

The reaction was incubated for 30 min at 37°C, 30 min at 60°C and 10 min at 80°C, sequentially.

One hundred ng of pACD4K-C plasmid DNA obtained from Sigma Aldrich company was also digested with *Hind* III and *Bsr* G I under conditions outlined in the table below.
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 µl</td>
<td>10X N.E. Buffer 2</td>
</tr>
<tr>
<td>5 µl</td>
<td>pACD4K-C plasmid DNA (100 ng)</td>
</tr>
<tr>
<td>1 µl</td>
<td>BsrG I</td>
</tr>
<tr>
<td>1 µl</td>
<td>Hind III</td>
</tr>
<tr>
<td>0.2 µl</td>
<td>BSA (100 µg/ml)</td>
</tr>
<tr>
<td>11 µl</td>
<td>Nuclease free water</td>
</tr>
<tr>
<td>20 µl</td>
<td>Final volume</td>
</tr>
</tbody>
</table>

The reaction was incubated for 30 min at 37°C, 30 min at 60°C and 10 min at 80°C, sequentially.

3 Construction of pACD4K-C- gene 14/ gene 19/ gene 126 vector

3.1 Ligation reactions

The R.E. digested pACD4K-C plasmid DNA was resolved on a 0.9% agarose gel (Joseph et al., 2001) and 6.9 kb fragment of a gel was isolated and purified by phenol:chloroform:isoamylalcohol and precipitation method (as described previously). Final purified DNA was resuspended in TE buffer (10 mM Tris and 1 mM EDTA at pH 8.0) to the final concentration of 20 ng/ml. Hind III and BsrG I digested PCR products was also purified by phenol:chloroform:isoamylalcohol and precipitation method. Subsequently, the PCR DNA and isolated pACD4K-C plasmid DNA segment were used in the ligation reaction outlined below.
<table>
<thead>
<tr>
<th>Volume</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 μl</td>
<td>pACD4K-C linear Vector (40 ng)</td>
</tr>
<tr>
<td>1 μl</td>
<td>BsrGI/Hind III-digested intron PCR product (10 ng)</td>
</tr>
<tr>
<td>1 μl</td>
<td>10X T4 DNA ligase buffer</td>
</tr>
<tr>
<td>1 μl</td>
<td>T4 DNA ligase (Promega, Madison)</td>
</tr>
<tr>
<td>5 μl</td>
<td>Nuclease free water</td>
</tr>
<tr>
<td>10 μl</td>
<td>Total volume</td>
</tr>
</tbody>
</table>

Ligate at 16 °C for overnight

3.2 Transformation in *E. coli* and selection of positive clones

Two hundred μl of chemical competent cells of *E. coli* are mixed with 50 μl of 100 mM CaCl₂ and 49 μl of sterile water. One μl of ligation products were added to this suspension and mixed well by gentle tapping at the bottom of the tube. The contents were incubated on ice for 15 min, followed by a heat shock at 42°C for 2 min. The cells were incubated at room temperature for 10 min, added 1 ml of LB medium, and incubated at 37°C in a shaker incubator for 1 h. Two hundred μl of culture were transferred onto LB agar plates containing 25 mg/ml chloramphenicol, and also including 30 μl each of 20 mg/ml X-Gal and 100 mM IPTG spreaded onto the plates. The plates were incubated overnight in a 37°C incubator. White colonies were randomly selected and screened for the presence of retargeted insert containing pACD4K-C plasmids.

3.3 Extraction of Plasmid DNA

Plasmid DNAs were extracted by boiling preparation method (Joseph *et al.*, 2001) from white colony cultures grown in LB medium with chloramphenicol (25 μg/ml). Briefly, 1.5 ml of overnight grown bacterial cultures was transferred into micro-centrifuge tubes
and centrifuged 12,000 g for 5 min. The supernatant was aspirated carefully and cell pellet was resuspended in 0.4 ml plasmid lysis buffer (10mM Tris-HCl pH 8.0, 0.1 M NaCl, 1mM EDTA, 5% v/v Triton X-100). Twenty five μl of freshly prepared lysozyme (10 mg/ml) in 10 mM Tris-HCl buffer (pH 8.0) was added. The contents of the tube were vortexed for few sec to mix, placed in a boiling water bath exactly for 40 sec, and centrifuged at 12,000 g for 15 min at 4°C. The pellet containing cell debris was removed with the help of a tooth pick. Four hundred and twenty μl of cold (-20°C) 100% isopropanol was added to the supernatant and mixed by vortexing, incubated at room temperature for 5 min and centrifuged at 12,000 g for 15 min to recover plasmid DNA. Supernatant was discarded and the DNA pellet was rinsed with 1 ml of 70% ethanol and dried in a speed-vac system (Labconco Centrivap Concentrator, Kansas City, MO) typically for about 5 min. Final pellet was resuspended in 100 μl of TE buffer and contaminating bacterial RNA was removed by treating with 1 μl of RNase A (1 mg/ml) at 37°C for 5 min. The quantity and quality of the plasmid DNA were checked by agarose gel electrophoresis.

3.4 Confirmation by Restriction Enzyme reaction

The extracted plasmids (5 μl each) were digested with BsrG I and Hind III enzyme similar to the method outlined previously. The clones with the correct inserts were selected for sequencing analysis to verify mutation in the modified group II intron segments.
3.5 Confirmation by Sequencing

The final constructs of pACD4K-C-Ech_1136 (p28-Omp gene 14), pACD4K-C-Ech_1143 (p28-Omp gene 19) and pACD4K-C-Ech_0126 (p28-Omp gene 126) (all three constructs are also called pACD4K-C-gene 14/gene 19/gene 126) were sequenced by gene analytic system, using T7 primer (forward primer) and one gene specific primer (IBS or EBS1d). The protocols for sequencing analysis were followed as described in the CEQ 2000 Dye Terminator Cycle Sequencing with Quick Start Kit manual (Beckman Coulter, CA). Sequences were manually evaluated to confirm mutations specific for targeting genes _1136 (gene 14), Ech_1143 (gene 19) and Ech_0126 (gene 126).

4 Cloning the tuf promoter into a promoterless vector, pBlue-TOPO

4.1 Cloning tuf promoter

RRG725 (tuf-Forward primer) and RRG726 (tuf-Reverse primer) were designed based on the E. chaffensis, Arkansas whole genome sequence information obtained from Genbank (accession CP000236). The primers were obtained from Intergrated DNA technologies (Coralville, Iowa). These primers and E. chaffeensis genomic DNA were used to amplify ‘tuf’ promoter region. Details of the PCR experiment were described below.
PCR reaction mixture

<table>
<thead>
<tr>
<th>Volume</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 ml</td>
<td>HiFi PCR buffer (10 X)</td>
</tr>
<tr>
<td>0.5 μl</td>
<td>d NTP (10 mM)</td>
</tr>
<tr>
<td>1 μl</td>
<td>MgSO₄ (50 mM)</td>
</tr>
<tr>
<td>0.5 μl</td>
<td>RRG 725 (10 μM)</td>
</tr>
<tr>
<td>0.5 μl</td>
<td>RRG 726 (10 μM)</td>
</tr>
<tr>
<td>0.2 μl</td>
<td>HiFi polymerase</td>
</tr>
<tr>
<td>1 μl</td>
<td>E. chaffeensis (7000) diluted</td>
</tr>
<tr>
<td>14.3 μl</td>
<td>Nuclease free water</td>
</tr>
<tr>
<td>20 μl</td>
<td>Total volume</td>
</tr>
</tbody>
</table>

Temperature cycles

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>94°C for 2 min</td>
</tr>
<tr>
<td>40 cycles</td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C for 30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>51°C for 30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td>68°C for 40 sec</td>
</tr>
<tr>
<td>Final extension</td>
<td>68°C for 10 min</td>
</tr>
<tr>
<td>Soak</td>
<td>4°C</td>
</tr>
</tbody>
</table>

Predicted size PCR products (0.37 kb) were confirmed after resolving on a 0.9% agarose gel. Subsequently, two μl of the PCR product was used to ligate into the linear pBlue TOPO vector and recombinant plasmids containing the inserts were selected by following protocols outlined by the manufacturer (Invitrogen, CA).
4.2 Evaluation of *tuf* promoter activity by β-galactosidase assay

β-galactosidase assay was performed by following instruction manual of β-galactosidase assay kit (Agilent Technologies, CA). Briefly, *E. coli* transformants containing the recombinant plasmids (pBlue-TOPO-*tuf*) were cultured overnight in 3 ml LB medium containing 50 µg/ml ampicillin, and centrifuged at 250 g for 5 min to remove the growth medium. One ml 1 X phosphate buffered saline (PBS) was added into each tube and centrifuged at 250 g for 5 min. The pellet was resuspended in 100 µl 1 X lysis buffer, and the sample was frozen on dry ice and thawed in a 37°C water bath to promote bacteria lysis. This step of freezing and thawing was repeated twice. Then, the sample was centrifuged at 12000 g for 5 min in 4°C, and the supernatant was transferred into a microtiter plate. Samples were diluted at 1:100 and 10 µl of each was transferred into the wells containing 50 µl 1 X cleavage buffer with β-mercaptoethanol and 17 µl o-Nitrophenyl-β-D-galactosidase (ONPG). The plates were incubated in 37°C for 30 min. The reaction was stopped by adding 125 µl stop buffer. The absorbance at 420 nM and the specific activity of *tuf* promoter (as assessed by the production of β-galactosidase per milligram of protein) were measured.

\[
N \text{ moles ONPG hydrolyzed} = \frac{(OD_{420}) (1.92 \times 105 \text{ nl})}{(4500 \text{ nl/ nmole-cm}) (1\text{cm})}
\]

Specific activity = n moles ONPG hydrolyzed / mg/ protein

5 Construction of pACD4K-C-gene 14/gene19/gene 126-*tuf* vector

PBlue-TOPO-*tuf* and pACD4K-C-gene 14/gene 19/gene 126 were digested with Hind III, the desired fragments were isolated, phenol purified and then ligated using T4 DNA ligase at 4°C overnight. One µl of the ligation mixture was transformed into XL1-Blue
strain of *E. coli* chemically prepared competent cells by heat shock method as described above. Transformed cultures were plated onto LB plate containing chloramphenicol (25 µg/ml) and incubated at 37°C overnight. The colonies were screened for the insertion and correct orientation by plasmid DNA isolation, followed by *Afl* II and *Hind* III digestion and by sequence analysis. Methods for restriction enzyme digestion, ligation, and cloning and sequence analysis were similar to the methods previously described in this chapter. The constructs with the correct inserts and orientations were saved in -80°C for future use.

**6 Construction of pACD4K-C-gene 14/19/126-tuf-rpsl-CAT**

The pACD4K-C-gene14/19/126-tuf plasmids were used to insert antibiotic cassette, chloramphenicol acetyltransferase (CAT) during expression using a transcriptionally active *E. chaffeensis* promoter, *rpsl*. The DNA segment of *rpsl*-CAT was already prepared and is available in our laboratory as a recombinant plasmid. Primers were designed to amplify *rpsl*-CAT DNA segment. Both the primers included a *Mlu*I site at the 5’ ends to facilitate cloning into Targetron constructs. PCR products with *Mlu*I digested fragment were prepared and used for ligation with *Mlu*I digested plasmids of pACD4K-C-gene14/19/126-tuf, T4 DNA ligations were performed as described previously. One µl of the ligation mixture was transformed into XL1-Blue chemical competent cells. The positive clones were screened and *rpsl*-CAT insert containing plasmids were verified by *Mlu*I digestion of plasmid DNA. The presence of inserts was also confirmed by plasmid DNA sequence analysis. The final constructs were used for preparing large quantity of plasmid DNAs by using Qiagen plasmid DNA maxi kit (Qiagen Inc. CA) and saved at -80°C for future use.
7 Evaluating the transcription from Group II intron in *Escherichia coli*

7.1 Performing RNA isolation from *E. coli*

PACD4K-C-gene14/19/126-tuf-rpsl-CAT plasmids culturing *E. coli* cultures were grown in 5 ml of LB media containing chlorepheicol (25 µg/ml) for almost 14 h at 37°C. Total RNA from the cultures were isolated using Tri-reagent method (Sigma Chemical Corporation, St. Louis, MO). Briefly, 1.5 ml of culture was harvested by centrifuging at 12,000 g for 5 min at 4°C. To the cell pellet, 1ml of tri-reagent was added and vortexed to mix. The lysates were incubated at room temperature for 5 min. Two hundred µl of chloroform was added to the cell suspension, mixed by vortexing and incubated at room temperature for 10 min. Subsequently, the samples were centrifuged at 12,000 x g for 15 min at 4°C. The top aqueous layer containing RNA was transferred into a clean micro centrifuge tube, mixed with 0.5 ml of cold isopropanol (-20°C) and incubated at room temperature for 10 min. The contents are centrifuged at 12,000 g for 15 min at 4°C. The RNA pellet recovered was rinsed with 70% ethanol and then resuspended in 100 µl of nuclease free water. To prevent the degradation of RNA, 40 µl of RNase inhibitor (Ambion Corporation, Austin, TX) was added and stored at -80°C until use.

7.2 RNA analysis

The total RNA was treated with RQ1 RNase-free DNase enzyme. Gel rig including combs and gloves were treated by washing with DEPC treated H2O. Agarose gel was made by mixing 0.6 g agarose in 28.8 ml water and microwaving for 30 sec. When it cooled down to 60°C, 8 ml of 36% formaldehyde and 4 ml of 10 X MOPS were added...
into agarose gel. RNA samples were prepared by adding 10 µl formaldehyde loading dye, heating at 65 °C for 15 min and chilling on ice. RNA running buffer was prepared by adding 25 µl of 10 X MOPS into 225 µl of DEPC. All of the samples were loaded onto gel and run at 70 V for 1 h.

7.3 RT-PCR analysis

Two PCR and one RT-PCR were performed using 10 µM Primers (IBS and EBS1d). RT-PCR for extracted RNA was designed to confirm the transcription of pACD4K-C-gene 14/19/126-tuf-rpsl-CAT in E. coli. PCR for extracted RNA was designed as the negative control for the transcription, and PCR for plasmid DNA was designed as the positive control for the transcription. In addition, another negative control was set up by using RNase free water as the templates. The RT-PCR reaction was performed as follows: reverse transcription at 48°C for 1 h; initiation at 94°C for 30 sec; and 25 cycles for (denaturation at 94°C for 15 sec, annealing at 55°C for 30 sec and extension at 72°C for 30 sec); then elongation at 72°C for 2 min, and final soaking at 4°C.

8 Preparation of Ehrlichia chaffeensis for transformation experiments

8.1 In vitro cultivation of E. chaffeensis:

The Arkansas isolate of E. chaffeensis was cultivated in vitro in the canine macrophage cell line DH82 at 37°C, by following the protocols established before (Chen et al., 1994; Munderloh et al., 1999).

8.2 Purification of E. chaffeensis:
*E. chaffeensis* infected macrophage cultures were harvested when the infectivity reached to 80-90% in nearly 100% confluent flasks. Infection was assessed by microscopic examination of polychromatic stained cytospin slides. About 25 ml of infected culture was collected into a 50 ml sterile falcon tube and centrifuged at 15,000 g for 15 min. The supernatant was discarded and the pellet is resuspended in 10 ml of 0.25 M ice cold sucrose solution. The cells were lysed by adding 1 mm diameter glass beads and vortexing twice for 30 sec. The lysed cell suspension was transferred to a new sterile tube free of glass beads and centrifuged at 100 g for 10 min. The supernatant was carefully collected and transferred to a new tube and centrifuged at 15,000 g for 15 min to collect cell-free *Ehrlichia* organism. At the end of centrifugation, supernatant was discarded and the cell-free bacteria pellet was resuspended in 10 ml of 0.25 M ice cold sucrose. The centrifugation step was repeated one more time and the final *E. chaffeensis* organism pellet was resuspended in 0.6-1 ml of 0.25 M sucrose solution.

### 8.3 Preparation of plasmid DNA for electroporation

Grow 100 ml of pACD4K-C gene 14/19/126-*tuf*-*rpsl*-CAT plasmids containing *E. coli* in LB media with chloramphenicol (25 µg/ml) for about 16 h. The cultures were used to isolate plasmid DNA by following the protocols outlined by EndoFree Plasmid Purification Kit (QIANGEN, CA). Briefly, the culture was centrifuged at 6,000 g for 15 min at 4°C, and then the supernatant was discarded. The pellet was resuspended in 10 ml buffer P1 (QIANGEN, CA), and mixed thoroughly with 10 ml buffer P2 (QIANGEN, CA), followed by incubating at room temperature 24°C for 5 min. During the incubation, the QIAfilter Cartridge was prepared (screw the cap onto the outlet nozzle of the
QIAfilter Maxi Cartridge and place the QIAfilter Cartridge in a convenient tube. Ten ml chilled buffer P3 (QIANGEN, CA) was added to the lysate, and mixed immediately and thoroughly by vigorously inverting for 6 times (do not incubate the lysate on ice). The lysate was poured into the barrel of the QIAfilter Cartridge, and incubated at room temperature for 10 min. The cap from the QIAfilter Cartridge outlet nozzle was removed. The plunger was gently inserted into the QIAfilter Maxi Cartridge and the cell lysate was filtered into a 50 ml tube. Two point five ml buffer ER (QIANGEN, CA) was added to the filtered lysate, mixed by inverting the tube approximately 10 times, and incubated on ice for 30 min. A QIAGEN-tip 500 was equilibrated by applying 10 ml buffer QBT (QIANGEN, CA), and allowing the column empty by gravity flow. The filtered lysate was applied to the QIAGEN-tip. The QIAGEN-tip was washed with 30 ml buffer QC (QIANGEN, CA) for 2 times. Plasmid DNA was eluted with 15 ml buffer QN (QIANGEN, CA) and precipitated by adding 10.5 ml room-temperature isopropanol, followed by mixing and centrifuging immediately at 15,000 g for 30 min at 4°C. The supernatant was carefully decanted. DNA pellet was washed with 5 ml of endotoxin-free room-temperature 70% ethanol (40 ml of 96-100% ethanol was added to the endotoxin-free water supplied with the kit) and centrifuged at 15,000 g for 10 min. The supernatant was carefully decanted without disturbing the pellet. The pellet was air-dried for 10 min, and the DNA was redissolved in 100 µl endotoxin-free buffer TE (QIANGEN, CA). The purified plasmid was stored at -80°C for later use.

9 Transformation with *E. chaffeensis* in DH 82 macrophage cells
E. chaffeensis cultures were grown in macrophage cell line, DH 82, and when infection reached to > 90%, cultures were harvested by spinning at 15,000 g for 15 min at 4°C. To the pellet, 10 ml of ice-cold 0.25 M sucrose solution was added and vortexed to resuspend the culture. The host cells were broken by passing through 27.5 gauge bended needles for 4 times, and spinning at 200 g for 10 min at 4°C to pellet the cell debris. The supernatant was filtered through 2.7 µm filter and spun at 15,000 g for 15 min at 4°C to collect the cell free bacteria. The pellet was washed twice with ice cold sucrose solution (0.25 M) and the final pellet was resuspended in 0.6-1 ml of 0.25 M ice-cold sucrose. The final sucrose solution varied depending on the bacteria volume and 100 µl each of the purified E. chaffeensis suspension was used for transformation expressions. One µl plasmid DNA (20 µg/ml) of pACD4K-C-gene 14/19/126-tuf-rpsl-CAT and 100 µl of cell free bacteria was mixed in an electroporator cuvette for use in electroporation. The electroporation was performed twice at 1800 V. To serve as a control, 100 µl of cell-free bacteria were also electroporated by giving double shock at 1800 V. The electroporated cells were transferred into T 25 cm² flask of DH82 cells having 80% confluency (5 ml culture volume). The electroporator cuvettes were washed with 400 µl of MEM medium, and then transferred the contents to the respective culture flask. The culture flasks were incubated at 37°C for overnight. Fraction of cultures from each flask (1.5 ml each) were harvested at different times post transformation (6 h, 12 h, 24 h, 48 h and 6 days) for use in DNA isolation and analysis. To the remaining culture, media is replaced every three or four days and E. chaffeensis infection was monitored for up to two months. The culture flasks with 2 µg/ml chloramphenicol were starting from the day six following transformation.
10 Methods for monitoring transformants

Total genomic DNAs of recovered cultures were isolated using QIAamp DNA Mini Kit (QIAGEN, CA). PCR analysis was performed by using primers spanning the expected insertion regions, one targeted to group II intron mutated genes and the other targeted to *E. chaffeensis* genome. The presence of genetically modified *E. chaffeensis* was also assessed by microscopic analyzing polychromatic stained culture slides and evaluating the growth or lack of growth of the mutants. Any culture that appeared to be positive was further evaluated by molecular techniques.

The PCR products were then resolved on a 0.9% agarose gel by subjecting to 40 V for 6 h and the resolved DNA was transferred to a nylon membrane. The prehybridization solution stock was prepared at final concentration of (6 X SSC, 10 mm pH 6.8 NaPO4 buffer, 1 mm EDTA pH 8.0, 10 X Denhard’s, 100 µg/ml ssDNA, and 0.5% SDS). The membrane was prehybridized for 2 h at 68 °C. The ³²P-labeled DNA probes for the group II intron was synthesized by using Primer-It II Random primer labeling kit (Stratagene, CA). After 14 h of hybridization at 68 °C, the membrane was washed once with 6 X SSC, 0.1% SDS followed by one wash at 2 X SSC, 0.1 % SDS, one wash at 1 X SSC, 0.1% SDS, and a final wash at 0.1 X SSC, 0.1% SDS. The nylon membrane was then exposed to an X-ray film at -70°C by using an intensifying screen. The X-ray film was developed and presence of group II intron segments within *E. chaffeensis* genome was assessed from the autoradiograph.
Reference:


Chapter III

Evaluation of Targetron Based Mutagensis in *Ehrlichia chaffeensis*
ABSTRACT

_Ehrlichia chaffeensis_ is an emerging tick-borne rickettsial pathogen that causes infection in people and several vertebrate animals. One of the striking features of _E. chaffeensis_ infection is the prolonged persistence in its vertebrate and tick hosts. The mechanisms of persistent infection and the reasons for the host immune system failure to clear the infection are not well understood. One hypothesis is that differential gene expression serves as an important adaptive mechanism used by _E. chaffeensis_ in support of its continued survival in both tick and vertebrate hosts. One way to test this hypothesis is by performing mutational analysis. The methods for introducing mutations in this pathogen have not yet been documented and are challenging, possibly due to its obligate, intraphagosomal growth requirement. A novel gene mutation method, called ‘TargeTron Gene Knockout System’ that is based on the modified group II intron insertion strategy has been developed recently, are used to create mutations in a wide range of gram positive and gram negative bacterial organisms. The objective of this study is to create mutations in _E. chaffeensis_ using the TargeTron gene knockout system. Three _E. chaffeensis_ genes: Ech_0126 (a transcriptionally silent gene), macrophage-specific expressed gene (p28-Omp 19) and tick cell-specific expressed gene (p28-Omp 14) were evaluated for mutational analysis. Group II intron was modified for retargeting to these genes and its expression in _E. chaffeensis_ is accomplished using the pathogen-specific high-expressing gene promoter (tuf) inserted upstream to the transcription starting site. In addition, a chloramphenicol acetyltransferase gene with _E. chaffeensis rpsl_ promoter (another transcriptionally active, high expression gene promoter) was introduced for use
as an antibiotic selection marker. The constructs were evaluated by transforming into *E. chaffeensis*. Transformants with mutations introduced in two of the three genes (Ech_0126 and p28-Omp 19) were identified by PCR and Southern blot methods for up to 48 h. This is the first report which demonstrates the creation of mutations in *E. chaffeensis*. Establishment of stable transformants, however, requires additional modifications to the constructs. The outcomes of this study will have important implications in defining the pathogenesis of *E. chaffeensis* and other closely related bacteria to assess the functional significance of host specific differential expression in the organisms in tick and vertebrate hosts.
**Introduction**

*E. chaffeensis*, an intracellular tick-transmitted rickettsial, is the agent of an infectious disease termed human monocytic ehrlichiosis (HME). Despite the increased incidence of HME in the USA (Paddock *et al.*, 2003; McNabb *et al.*, 2008), efficient methods to control the illness remain to be established.

Recent studies demonstrate that the pathogen gene expression is significantly different when it replicates in tick cell compared to its growth in vertebrate macrophages (Ohusi *et al.*, 1998; Unver *et al.*, 2002; Long *et al.*, 2002; Singu *et al.*, 2005; Seo *et al.*, 2008). Studies which evaluated host immune response in the mouse model also suggested that the tick cell-derived *E. chaffeensis* persist pathogen (Ganta *et al.*, 2002; Ganta *et al.*, 2004; Ganta *et al.*, 2007). The delayed clearance in the murine host is also associated with altered host immune response (Ganta *et al.*, 2002; Ganta *et al.*, 2004; Ganta *et al.*, 2007). The evaluation of differential expressed *E. chaffeensis* proteins will be useful in creating mutation.

‘TargeTron Gene Knockout System’ that is based on the modified group II intron insertion strategy has been utilized for creating mutations in several gram negative and gram positive bacteria. In this study, we evaluated the feasibility of creating mutations in *E. chaffeensis* using Group II intron based strategy. Three *E. chaffeensis* genes were assessed and the results from this study and their implication were discussed.
Result

1 Construction of plasmids for retargeting group II intron for inserting mutations in *E. chaffeensis* genome

1.1 PCR for intron re-targeting in genes of *E. chaffeensis*:

Genomic sequences of *E. chaffeensis* genes, Ech_1136 (p28-Omp gene 14), Ech_1143 (p28-Omp gene 19) and Ech_0126 were uploaded to Sigma-Aldrich company website ([www.sigma-aldrich.com/targetron](http://www.sigma-aldrich.com/targetron)), and three primers (IBS, EBS-1d, EBS-2) for each gene were designed and synthesized commercially for each gene-specific retargeting (Table 3.1) (Gene-specific retargeting regions were identified in italics and bold text in the primers). The PCR amplification was performed using group II intron as template using each gene-specific retargeted primer. Retargeted PCR products of 0.35 kb for each gene were identified in the ethidium bromide stained agarose gel (Figure 3.1).

Figure 3.1  PCR for intron retargeting in gene 14, gene 19 and gene 126 of *E. chaffeensis*. Lanes 1, 2, 3 are Ech_1136 (p28-Omp gene 14), Ech_1143 (p28-Omp gene 19) and Ech_0126, respectively.
1.2  Ligation and recombination plasmid construction

1.2.1 Restriction enzyme digestion

To engineer the retargeted PCR products into group II intron cassette, PCR products and pACD4K-C plasmids (Sigma, St. Louis, MO) were digested separately with *Hind* III and *BsrG* I restriction enzymes. The plasmid segment of 6.9 kb containing the active group II intron fragment, but lacking the *lacZα* segment of 0.76 kb was isolated and ligated with RE.digested retargeted PCR products for genes Ech_1136, Ech_1143 and Ech_0126 using T4 DNA ligase. The ligated products were transformed into *E.coli* strain, XL1 Blue, plated on agar plates containing chloramphenicol (25 µg/ml), white colonies were picked and screened for the presence of the retargeted inserts.

![Figure 3.2 Double digestion of the retargeted PCR products and pACD4K-C plasmids with *Hind* III and *BsrG* I. Lanes 1 and 2 are pACD4K-C plasmids; lanes 3, 4, and 5 are retargeted PCR product for Ech_1136, Ech_1143, and Ech_0126; M represents marker in base pairs.](image)
1.22 Confirmation of constructed plasmids pACD4K-C-gene 14/19/126 by restriction enzyme digestion. The plasmid DNAs were isolated from the liquid cultures of transformants by boiling preparation method and assessed for the presence of specific inserts size (Figure 3.3). Fragments of about 350 bp Hind III and BsrG I digested products were identified in several clones.

![Image of gel electrophoresis](Image)

**Figure 3.3** Double digestion of constructed plasmids pACD4K-C-gene 14/19/126 with Hind III and BsrG I (the arrow indicates the 350 bp size fragments).

1.3 Verification of modified intron

Three clones from each gene were randomly selected and the sequence modifications in pACD4K-C-gene 14/gene 19/gene 126 plasmids were determined by CEQ-8000 gene analytic system for DNA sequencing. These modified plasmid containing *E. coli* cultures were stored in -80°C as glycerol stocks for subsequent use in isolating plasmid DNAs.

1.4 Insertion of *E. chaffeensis* promoter
To drive the expression of modified group II introns in *E. chaffeensis*, we identified a strong transcriptionally active *E. chaffeensis* gene promoter, *tuf*. This promoter activity was primarily determined by our group (Unpublished data). To clone the promoter region of *tuf*, two primers were designed (RRG 725 and RRG 726) (Table 3.1) and used in a PCR to amplify from *E. chaffeensis* genome. To facilitate cloning upstream to modified group II introns in pACD4K-C plasmids, *Hind* III restriction sites sequence were engineered in the primers. PCR products were evaluated in 0.9% agarose gel by subjecting to electrophoresis at 120 V for 1 h. Predicted fragment of 374 bp was detected in the gel under UV light (Figure 3.4).

![Figure 3.4  PCR for tuf promoter.](image)

*E. chaffeensis* *tuf* promoter PCR products were cloned into pBlue-TOPO vector by following the protocol described by Invitrogen pBlue-TOPO TA Expression Kit (Invitrogen, CA). The pBlue-TOPO plasmid contains a promoterless *LacZ* gene placed down stream to multiple cloning site regions. If the promoter inserted (such as the *tuf*) is active, β-galactosidase proteins will be made in *E.coli* cultures containing the
recombinant plasmid. Tuf promoter inserted pBlue-TOPO plasmids were assessed for transcriptional activity by β-galactosidase assay. The results showed that the transcriptional activities of plasmids containing tuf promoter are 11.58 and 15.3 times higher (Figure 3.5; lanes 2 and 3) than the control (pBlue-TOPO vector without inserts, lane 1), suggesting that tuf promoter of E. chaffeensis is a strong promoter during the expression of β-galactosidase in E. coli.

![beta-gal activity of tuf promoter](image)

**Figure 3.5** β-galactosidase activity of tuf promoter. Lane 1, lysates of E.coli containing pBlue-TOPO plasmid without inserts; lanes 2 and 3 contained E. coli lysates prepared from two randomly selected clones of pBlue-TOPO vector with tuf promoter.

1.4.1 Restriction enzyme digestion

The plasmids pACD4K-C-gene 14/19/126 and a tuf promoter plasmid of pBlue-TOPO were digested by Hind III and digested products were isolated and analyzed in 0.9 % agarose gel to verify the recovery of DNA fragments. Fragments of 7,677 bp of
pACD4K-C- gene 14/19/126 linear DNA and 374 bp of tuf promoter were isolated and observed under UV light (Figure 3.6).

1.4.2 Cloning of tuf promoter in front of group II introns in pACD4K-C plasmids

Purified Hind III digested pACD4K-C plasmid segments with retargeted group II introns for E. chaffeensis genes for Ech_1136, Ech_1143, and Ech_0126, and tuf promoter DNA were ligated, cloned into E. coli (TOP 10 strain), and plasmids with correct orientation were selected by following standard molecular cloning methods. The constructed plasmids pACD4K-C-tuf were picked from bacteria culture plates containing chlorehpenicol (25 µg/ml) and cultured overnight in 3 ml of LB plus chlorehpenicol (25 µg/ml). The plasmid DNAs from liquid cultures were extracted by boiling preparation method and digested with Afl II and EcoRV to confirm the orientation of tuf promoter inserts. For gene 19 and gene 126, fragments of 7,219 bp plus 814 bp should be
observable in the plasmids with forward insertion (and a 18 bp fragment should also be released); and fragments of 6849 bp plus 814 bp and 388 bp should be observable in the plasmids with reverse insertion. For gene 14, fragments of 7219 bp plus 832 bp should be observable in the plasmids with forward insertion; and fragments of 6849 bp plus 1202 bp should be observable in the plasmids with reverse insertion. The result indicate that both gene 14 plasmids are in forward orientation; one each of the gene 19 plasmids are in forward and reverse orientation (lanes 3 and 4); and one of the gene 126 plasmids has forward orientation (lane 6) and the other does not have any inserts (lane 5).

Figure 3.7  Confirmation of pACD4K-C-gene 14/19/126-tuf. Lanes 1 and 2, Ech_1136; lanes 3 and 4, Ech_1143; and lanes 5 and 6, Ech_0126.

The final constructs of pACD4K-C-gene 14/19/126-tuf were verified by sequencing analysis in CEQ-8000 gene analytic system. The verified constructs containing *E. coli* cultures in TOP 10 strain were stored as glycerol stocks at -80°C for long term use.

1.5 Insertion of *rpsl*-CAT
1.5.1 PCR for rpsl-CAT

Rpsl-CAT fragment was amplified by PCR, using previously prepared plasmid construct containing the *E. chaffeensis* rpsl promoter and CAT gene coding sequence (available in our laboratory). After PCR reaction, 3 µl of PCR products were analyzed in an agarose gel to identify predicted fragment of 948 bp.

![Figure 3.8](image.png)

**Figure 3.8** PCR for rpsl-CAT.

1.5.2 Restriction enzyme digestion of the PCR products of rpsl-CAT and the plasmids of pACD4K-C-gene 14/19/126-tuf. PACD4K-C constructs contain a kanamycin gene cassette downstream to group II intron and within the intron encoding protein coding region. This cassette can be replaced with an alternate antibiotic selection marker, such as chloramphenicol acetyltransferase (CAT), as it is present between two *Mlu*I restriction enzyme sites as a 1.37 kb insert. The PCR products of rpsl-CAT (0.95 kb) were also engineered to contain *Mlu*I sites at both ends of the segment. The PCR products of rpsl-CAT and the plasmids of pACD4K-C-gene 14/19/126-tuf were digested by *Mlu*I. Digested products were resolved in an agarose gel to identify predicted fragments (6.67 kb and 1.37 kb from plasmids and 0.95 kb for rpsl-CAT (Figure 3.9)).
Figure 3.9 Double digestion with *Mlu*I. Lanes 1, 2, 3 and 4 are digested PCR products of *rpsl*-CAT, the digested plasmid of pACD4K-C-gene 14, the digested plasmid of pACD4K-C-gene 19, and the digested plasmid of pACD4K-C-gene 126, respectively.

The *Mlu*I digested fragment of 6.67 kb for each of the three plasmid constructs for Ech_1136, Ech_1143, and Ech_0126 genes and 0.95 kb of *rpsl*-CAT were isolated and used in ligation and cloning experiments in *E. coli* (TOPO 10 strain) to recover plasmids containing the *rpsl*-CAT inserts (Figure 3.10). The constructed plasmids pACD4K-C-gene 14/19/126-*tuf*-rpsl-CAT were prepared by boiling preparation method. The presence of *rpsl*-CAT inserts of 0.95 kb fragments were confirmed by digesting with *Mlu*I enzyme. The presence of correct sequence in the *rpsl*-CAT inserts was further confirmed by sequence analysis. The final constructs were stored in -80°C for later use.
2. Evaluating the transcription from Group II intron in *Escherichia coli*

PACD4K-C-gene 14/19/126-*tuf*-rpsl-CAT plasmids containing *E. coli* cultures were grown in 5 ml of LB media containing chlorephenicol (25 µg/ml) for almost 14 h at 37°C. Total RNA from the cultures was isolated using Tri-reagent method (Sigma Chemical Corporation, St. Louis, MO) (Figure 3.11). The total RNA was treated with RQ1 Rnase-free Dnase enzyme. RNAs were analyzed for the presence of transcripts of group II intron driven by *E. chaffeensis* *tuf* promoter. Specifically, a primer set targeted to retargeted segments (Forward primers are IBS and reverse primers are EBS1d) was used to assess the expression (Figure 3.12). No products were detected when no template was used or when RNA template was used in a PCR for the RNA from retargeted plasmid containing *E. coli* of gene 14 and gene 19. Weak positive product was observed in PCR product of gene 126. When RNA was used in RT-PCR, all three RNAs yielded the predicted amplicons. For gene 126 construct RNA, the amplified product for RNA is significantly more than that observed in PCR. This suggests that the presence of positive PCR product in gene 126 construct RNA in the absence of reverse transcriptase may be due to residual plasmid DNA contamination. Together, these results demonstrate that the
group II intron constructs prepared for genes 14, 19, and 126 are functional in making transcripts.

Figure 3.11 RNA gel of plasmids pACD4K-C-gene 14/19/126-tuf-rpsl-CAT. Lanes 1, 2, 3, 4, 5, 6 and 7 are the total RNA isolated using Tri-reagent methods.

Figure 3.12 RT-PCR, PCR for extracted RNA and PCR for plasmid DNA. Lanes 1, 5, and 9 are negative controls using Rnase free water as the template; lanes 2, 6, and 10 are PCR for plasmid DNA; lanes 3, 7 and 11 are PCR for extracted RNA; lanes 4, 8 and 12 are RT-PCR for extracted RNA.
3 Transformation and mutation analysis

The modified Targetron plasmids for use in creating mutations in *E. chaffeensis* genes, Ech_1136, Ech_1143, and Ech_0126 were used in transformation experiments. Typically, 20 µg of purified plasmid DNA was electroporated into purified *E. chaffeensis* organism recovered from about 10 ml of 80-90% infected macrophage cultures representing ~80% confluency. The experiments of transformation were performed three times. *E. chaffeensis* was subjected to electroporation with the plasmids pACD4K-C-gene 14/19/126-tuf-rpsl-CAT by double shock at 1800 V. The experiment also included a control electroporation reaction that contained only *E. chaffeensis* organism and a second control that contains *E. chaffeensis* organism that were not subjected to electroporation. Electroporated and non electroporated *E. chaffeensis* organism were cultured in DH82 macrophage cells and harvested at different times. In the first experiment, the cells were harvested at 6 h, 12 h and 24 h; in the second and third experiment, the harvested cultures included at 6 h, 24 h, 48 h and 6 days. The harvested cultures were used to recover total genomic DNA for assessing the integration of group II introns into *E. chaffeensis* genome at the predicted locations of Ech_1136, Ech_1143, and Ech_0126 genes (described below). After harvesting the last batch of culture for DNA isolation, chloramphenicol was added to the culture media at a concentration of 2 µg/ml. Infection in DH82 cultures were monitored for up to two months by microscopy analysis. To date, no positive infection was found from all three transformation experiments. For further testing the insertion, PCR and Southern blot were performed.

3.1 Detection of insertions assayed by PCR methods
Two types of PCRs were performed to assess the integration of group II introns into *E. chaffeensis* genome. In the first set of PCRs, primers were designed to bind to the region upstream and downstream to integration site (these primers were expected to yield positive PCR products independent of the insertions). In the second set of PCRs, one primer is targeted *E. chaffeensis* genome either upstream or downstream to the insertion sites. The second primer of the second set was targeted to group II intron segment. The strategy and expected amplified fragments were described in Figure 3.13. To assess the mutations, the PCR products were analyzed in a 0.9 % agarose gel as well as Southern blot analysis after transferring to a nylon membrane and hybridization with group III intron specific $^{32}$P-labeled DNA probes.
Figure 3.13  PCR strategies for detecting mutants.
As outlined in the cartoon of Figure 3.13, predicted fragments from various PCR samples were analyzed for agarose gels or in the Southern blot results from PCR DNA hybridized with group II intron probe. Figure 3.14 has the agarose gel data for first experiment. In non-transformants and transformants for gene 14, we expected to see 0.86 kb and 2.92 kb amplicons, respectively, when PCR I is performed. In the gel, we did detect 0.86 kb, but not the 2.92 kb (Lanes 2, 4 and 6 of Figure 3.14). Predicted 0.82 kb fragment for PCR IIa was also not visible in the analyzed PCR products of gene 14 transformants (Lanes 1, 3, and 5 of Figure 3.14). Similar experiments for gene 19 yielded predicted fragments for PCR I (0.24 kb for gene 19; lanes 8, 10 and 12 of Figure 3.14) and for gene 126 as well (2.7 kb; lanes 14, 16 and 18 of Figure 3.14). As observed from results of gene 14, expected amplicons for transformants for genes 19 and 126 (2.3 kb and 4.7 kb, respectively) were also not clearly visible. The analysis was also complicated with the presence of multiple non-specific amplified DNA fragments observed in the agarose gel. PCR IIa strategy should yield positive bands only if the genomic DNA included insertion mutations (for gene 14, 0.82 kb; for gene 19, 0.39 kb and for gene 126, 2.8 kb). The predicted fragments were undetectable for gene 14 PCR IIa products (Lanes 1, 3, and 5 of Figure 3.14). However, we do detect minor predicted bands for gene 19 and 126 PCR IIa amplicons. In addition, we observed multiple non-specific amplicons in the lanes representing PCR IIa of genes 19 and 126 (Lanes 7, 9 and 11 of Figure 3.14 for gene 19 and lanes 13, 15 and 17 of Figure 3.14 for gene 126).
Figure 3.14  Experiment 1, PCR assay for detecting insertion. Lanes 1, 2, 3, 4, 5 and 6 are gene 14; lanes 8, 9, 10, 11 and 12 are gene 19; and lanes 13, 14, 15, 16, 17 and 18 are gene 126; Lanes 1, 3, 5, 7, 9, 11, 13, 15, 17 are PCR IIa, and Lanes 2, 4, 6, 8, 10, 12, 14, 16, 18 are PCR I; Lanes 1, 2, 7, 8, 13, 14 are the electroporated cells harvested after 6 h; lanes 3, 4, 9, 10, 15, 16 are the electroporated cells harvested after 12 h; lanes 5, 6, 11, 12, 17, 18 are the electroporated cells harvested after 24 h.

Figure 3.15  Experiment 2, PCR assay for detecting insertion. Gene 14, gene 19 and gene 126 all have equal 6 lanes. For each of the six lanes, the templates for PCR assay are: from right to left, post 6 h harvested culture, post 24 h harvested culture, post 48 h harvested culture, post 6 days harvested culture, E. chaffeensis genome and nuclease free water.
Figure 3.16  Experiment 3, PCR assay for detecting insertion. Lanes 1, 2, 3, 4, 5 and 6 are PCR IIb for gene 19; lanes 7, 8, 9, 10, 11 and 12 are PCR IIa for gene 126; lanes 13, 14, 15, 16, 17 and 18 are PCR IIa for gene 19; lanes 19, 20, 21, 22, 23 and 24 are PCR I for gene 126; lanes 25, 26, 27, 28, 29 and 30 are PCR I for gene 19; and lanes 31, 32, 33, 34, 35 and 36 are PCR IIb for gene 126. For each of the six lanes, the templates for PCR assay are: from right to left, post 6 h harvested culture, post 24 h harvested culture, post 48 h harvested culture, post 6 days harvested culture, *E. chaffeensis* genome and nuclease free water.

### 3.2 Southern blot analysis

To further evaluate the insertion mutations, PCR DNAs resolved in the agarose gel were transferred to a nylon membrane and assessed by Southern blot using group II intron specific $^{32}$P-labeled DNA probe. In the first experiment, predicted DNA amplicons in PCR IIa products were detected only for gene 19 and 126, and no detectable signals were observed for gene 14 (Figure 3.17). The signal intensity for gene 19 was the highest for 12 h post-transformation and 6 h post-transformation for gene 126. For gene 126, we also
observed smaller size non-specific amplicons. The signal intensity decreased in the samples analyzed for 24 h of gene 19 and 12 h and 24 h for gene 126. In the experiment 2 and 3, group II insertion strategy, using the modified Targetron plasmids for *E. chaffeensis* genes 14, 19 and 126, was repeated two additional times. In the second experiment (Figure 3.18), only PCR IIa strategy was followed to assess the presence of mutations in *E. chaffeensis* genome. Similar to the first experiment, agarose gel analysis had no detectable amplicons for gene 14 and multiple bands for gene 19 and 126 transformants. The resolved DNAs were transferred to nylon membrane and hybridized with group II intron probe to see if we can identify predicted amplicons. Consistent with the first experiment, Southern blot data was negative for gene 14, and positive for genes 19 and 126. Insertion-specific amplicons decreased for 6 h post-transformation to 48 h post-transformation for gene 19 and undetectable in the DNA harvested after 6 days post-transformation. Similar analysis for gene 126 yielded positive only for 6 h time point. In the third experiment, we performed PCRs using the strategy of I, IIa, and IIb (as outlined in Figure 3.13). In addition, using the amplicons from the PCRs, we performed nested PCRs (two pairs of PCR primers were used. The first pair of primers amplified the locus as seen in any PCR experiments, the second pair of primers bind within the first PCR product and produce a second PCR product that will be shorter than the first one.). The aim for this experiment is to see if we can reduce the non-specific amplicons and also to enhance the signal intensity of the amplified products. The analysis for this experiment was performed for 6 h, 24 h, 48 h and 6 d similar to the second experiment.
Figure 3.17  Experiment 1, Southern blot for detecting insertion. 6 h: Southern blot for post 6 h harvested culture; 12 h: Southern blot for post 12 h harvested culture; and 24 h: Southern blot for post 24 h harvested culture.

Figure 3.18  Southern blot for experiment 2. 6 h: Southern blot for post 6 h harvested culture; 24 h: Southern blot for post 24 h harvested culture; 48 h: Southern blot for post 48 h harvested culture; and 6 D: Southern blot for post 6 days harvested culture.
Figure 3.19  Southern blot for experiment 3. Lanes 1, 2, 3, 4 are 6 h, 24 h, 48 h and 6 d post transformation culture of gene 19 using PCR IIa; lanes 5 is 6 h post transformation culture of gene 126 using PCR IIa.
Discussion

*E. chaffeensis* is an emerging tick-borne pathogen responsible for an important disease in people, human monocyte ehrlichiosis (Dawson *et al*., 1991; Lockhart *et al*., 1995). The disease was first discovered in 1987 and subsequently reported in many parts of the USA and world (Bakken *et al*., 1996; Bakken *et al*., 1998; Dumler *et al*., 1998; Whitlock *et al*., 2000; Standaert *et al*., 2000; Olano *et al*., 2003; Paddock *et al*., 2003; McNabb *et al*., 2008). *E. chaffeensis* also causes infections in several other vertebrate hosts (Dugan *et al*., 2000; Yabsley *et al*., 2004; Dugan *et al*., 2005; Yu *et al*., 2008). Despite recent progress, the pathogenesis and how the organism persists in vertebrate hosts and ticks still remain to be known.

Recent studies demonstrate differential expression of *E. chaffeensis* proteins in a host cell specific manner (Ohusi *et al*., 1998; Unver *et al*., 2002; Long *et al*., 2002; Singu *et al*., 2005; Seo *et al*., 2008). The protein expression differences are also shown to be influencing the host response and the pathogen persistence (Brelun *et al*., 1996; Cotter *et al*., 2000). In this study, we evaluated a method to create mutations in *E. chaffeensis* genome. Targetron mutagenesis system can be exploited to introduce mutations in virtually any bacterial organism (Yao *et al*., 2006; Cui *et al*., 2007; Heap *et al*., 2007; Rodriguez et al., 2008). However, to adapt to a species, such as to *E. chaffeensis*, one needs to make several manipulations to the Targetron constructs prior to use in creating mutations in the organism. First, we needed to identify suitable regions to create mutations in the *E. chaffeensis* genome. Second, we had to adapt the group II intron expression system for *E. chaffeensis* using a constitutively expressing gene promoter of
this organism. Third, we were required to introduce an antibiotic selection marker cassette into the Targetron constructs. The antibiotic resistance protein expression also requires an *E. chaffeensis* promoter that is stably expressed. Once the constructs were prepared, appropriate conditions for introducing foreign DNA into *E. chaffeensis* and methods to recover viable cultures were also necessary. In this project, our efforts were focused to develop all the above described methods.

We selected three genes of *E. chaffeensis* for creating mutations: 1) a gene which appeared to be transcriptionally silent (Ech_0126); 2) a gene predominantly expressing *E. chaffeensis* when it replicates in tick cells (Ech_1136; p28-Omp gene 14); and 3) a gene primarily expressed in the organism when grown in macrophages (Ech_1143; p28-Omp gene 19). The justification for selecting these genes are as follows: if a mutation is detrimental for *E. chaffeensis*, we may not be able to recover viable cultures following the introduction of such mutations. Therefore, we selected a transcriptally silent gene (Ech_0126) for one of the experiment. Assuming that this gene is not important for *E. chaffeensis* survival, we may be able to recover positive mutants if introduced in this gene. The choice of tick cell specific gene expression may allow the mutant *E. chaffeensis* to grow well in macrophage cultures, but may not allow its growth in tick cell cultures. Thus, we selected Ech_1136 for this type of mutations. This gene, also referred to as p28-Omp 14 gene, is primarily expressed in the organism replicating in ticks cells or in infected ticks (Ganta *et al.*, 2002; Ganta *et al.*, 2004; Ganta *et al.*, 2007). In the third selection, Ech_1143 was selected as it is abundantly expressed in macrophage cultures or infected vertebrate hosts (Ganta *et al.*, 2002; Ganta *et al.*, 2004; Ganta *et al.*, 2007). This
gene is also expressed at low levels in tick cell-derived bacteria (Ganta \textit{et al.}, 2002; Ganta \textit{et al.}, 2004; Ganta \textit{et al.}, 2007). This gene also belongs to the p28-Omp multigene locus and is referred to as p28-Omp 19. If expression from this gene is important for the bacteria growth in both host cell backgrounds or only critical in one or the other, we may see selective growth of mutants.

We selected an \textit{E. chaffeensis} “\textit{tuf}” promoter to drive the expression of modified group II introns targeted to Ech genes (Ech\_0126, Ech\_1136, and Ech\_1143). The \textit{tuf} promoter transcription was previously verified by our research team. This is a strong \textit{E. chaffeensis} promoter which is responsible for the transcription of more than 22 genes, all of which are housekeeping genes (GeneBank cp 000236.1). In this study, we confirmed the promoter activity of \textit{tuf} promoter in \textit{E.coli} for driving expression of β-galactosidase gene activity (Figure 3.5). Subsequently, the promoter was transferred to Targetron constructs. The RNA expression from \textit{tuf} promoter driving the expression of group II modified intron and intron encoded protein transcripts were also demonstrated in this study (Figure 3.11). These results were important prior to investigating the Targetron constructs in creating mutations in \textit{E. chaffeensis}.

In support of identifying \textit{E. chaffeensis} mutants, we inserted chloramphenicol acetyltransferase (CAT) gene coding sequence of \textit{E. coli} with another constitutively expressed \textit{E. chaffeensis} promoter, \textit{rpsl}. The \textit{rpsl} promoter belongs to protein synthesis complex genes, ribosomal proteins S12 (Ech\_0963). The CAT gene expression from \textit{rpsl} promoter was also confirmed in \textit{E. coli} by selecting clones resistant to chloramphenicol.
The transformant methods for *E. chaffeensis* required high volumes of purified, viable organisms. Our standardized protocols were effective in recovering viable *Ehrlichia* organism. In all three transformation experiments, we were able to demonstrate the growth of *E. chaffeensis* following electroporation. We also standardized conditions for transforming *E. chaffeensis* with Targetron constructs. Several electroporation conditions were tested which varied from 1200 V to 2400 V and single and double shocks. We found the most effective method of transformation is by providing two electric shocks at 1800 V.

Several pilot experiments were also performed to purify high quality DNA for use in transformation experiments. We found the Qiagen plasmid Endofree maxi kit (Qiagen, CA) to be the best in isolating high quality DNA. We also evaluated several plasmid DNA concentrations, which ranged from 2 μg to 20 μg and found the higher concentration of 20 μg to be the best for transformations. The selection 20 μg was based on the published evidence of transformations in other Rickettsial agents, *Rickettsia prowazekii* (Liu et al., 2007). In our final experiment of transformation, we were able demonstrate the integration of modified group II introns into *E. chaffeensis* genome at two locations: Ech_0126 and Ech_1143. Insertions at Ech_0126 (gene 126) were following for up to 6 days in all three experiments by PCR and Southern blot analysis prior to adding chloramphenicol. In the first experiment, the analysis was performed only for two days. Predicted (0.89 kb) amplicons were clearly visible following Southern blot analysis. Minor bands of expected size were also seen in the agarose gel resolved DNA of amplicons obtained for PCR IIa, in which one primer targeted to group II intron and the
second primer targeted to the genomic regions 5’ to predicted insertion site. The PCR positives in the first experiment, however, decreased with increasing time following transformation. The presence of PCR positives, however, was only seen for the first harvest, i.e. after 6 h post transformation in the cultured organisms in the second experiment. Similar transformation experiments for Ech_1143 (gene 19) also led to positive insertions in the genome and are detectable for up to 24 h in culture following the transformation. However, we found no evidence of mutations for gene 14 (Ech_1136). The detection of PCR positives for mutations in genes 126 and 19 are very encouraging and clearly demonstrate that *E. chaffeensis* genome can be mutated using Targetron-based insertion mutation methods. Introduction of mutations in *E. chaffeensis* or other species of *Ehlichia* and *Anaplasma* genome were not demonstrated earlier. Therefore, this study represents an important milestone in advancing the gap of knowledge in understanding the significance of host-specific protein expression.

It is not clear why mutated *E. chaffeensis* for gene 19 and 126 were not stable beyond 24 h. It is also not clear why the intensity of PCR positive bands decreased with time post transformation. One possible explanation for the unstable mutations in culture is that the mutants of genes 19 and 126 may be slow in their growth compared to wild type *E. chaffeensis*. The absence of positive mutants for gene 14 following Targetron-based mutagenesis analysis may be an indication that the loss of this gene function may lead to lethal outcome in *E. chaffeensis*. 
In our studies, we followed the Targetron mutated \textit{E. chaffeensis} of genes 19 and 126 for up to 2 months by adding the chloramphenicol at a concentration 2 µg/ml to the culture media. Our efforts to identify mutants by culture PCR did not result in positive clones. This result may be explained due to poor or no expression of chloramphenicol acyltransferase protein in \textit{E. chaffeensis}. This may be important, due to the difference in the codons of the CAT gene coding sequences. In particular, the CAT gene construct used in our studies was derived from \textit{E. coli}. The codon usage of \textit{E. chaffeensis} is significantly different from that of \textit{E.coli}. Future experiments, therefore, should involve optimizing the codons of CAT gene to \textit{E. chaffeensis}. The second possibility is the concentration of chloramphenicol. We may consider reducing the chloramphenicol to minimum required concentration to inhibit the growth of only the wild type \textit{E. chaffeensis}, but not the mutants. It is also important to investigate the possibility to create Targetron based mutations in other regions of \textit{E. chaffeensis} genome. All these experiment are now underway in our laboratory.

In summary, considerable progress has been made in introducing mutation in \textit{E. chaffeensis}. We are the first to demonstrate the feasibility to introducing mutations at two regions of \textit{E. chaffeensis} genome. We standardized numerous protocols needed in support of transformation and to create mutations in \textit{E. chaffeensis}. The Targetron constructs developed in this study can be adapted easily to target other genomic regions of \textit{E. chaffeensis} and also for use in creating mutations in other tick borne rickettsial agents of the genera \textit{Ehrlichia} and \textit{Anaplasma}. 
Reference:


Chapte IV

Major conclusions
1. Targetron modified constructs for *E. chaffeensis* genes Ech_1136, Ech_1143, and Ech_0126 were prepared.

2. *E. chaffeensis* “*tuf*” promoter was cloned and confirmed its function in *E. coli* using promoterless β-galactosidase expression system.

3. The *E. chaffeensis* *tuf* promoter driven expression of group II intron and intron encoded protein transcripts were also confirmed in *E. coli*.

4. The *E. chaffeensis* *rpsl* promoter was also cloned and utilized in driving an antibiotic selection marker in *E. chaffeensis*.

5. The final modified Targetron constructs include *E. chaffeensis* *tuf* promoter to drive the expression of modified group II introns and *rpsl* promoter to drive the expression of chloramphenicol acetyltransferase gene.

6. The Targetron constructs for *E. chaffeensis* genes 14, 19, and 126 were examined for transformation and creating mutations in the pathogen.

7. Mutants for *E. chaffeensis* genes 19 and 126 were identified in cultures for up to 48 h in three independent times.

8. No mutants were detected for gene 14.
9. The modified Targetron constructs are useful in preparing additional constructs for making mutations in alter regions of *E. chaffeensis* genome as well as for studying mutational analysis of other tick-borne pathogens of the genera *Ehrlicha* and *Anaplasma*.