Phylogenetic and Phylogeographic Studies of *Panstrongylus sp.*, Vectors of Chagas Disease in Loja and Manabí Provinces, Ecuador

A dissertation presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Carolina Sempertegui-Sosa

August 2012

© 2012 Carolina Sempertegui-Sosa. All Rights Reserved.
This dissertation titled

Phylogenetic and Phylogeographic Studies of *Panstrongylus sp.*, Vectors of Chagas Disease in Loja and Manabí Provinces, Ecuador

by

CAROLINA SEMPETEGUI-SOSA

has been approved for

the Program of Molecular and Cellular Biology

and the College of Arts and Sciences by

________________________

Mario Grijalva

Associate Professor of Biomedical Sciences

________________________

Howard Dewald

Interim Dean, College of Arts and Sciences
ABSTRACT

SEMPERTEGUI-SOSA, CAROLINA, Ph.D., August 2012, Molecular and Cellular Biology

Phylogenetic and Phylogeographic Studies of *Panstrongylus* sp., Vectors of Chagas Disease in Loja and Manabí Provinces, Ecuador  (171 pp)

Director of Dissertation: Mario Grijalva

The species *Panstrongylus howardi*, *P. chinai*, and *P. rufotuberculatus* were found infesting domestic, peri-domestic and sylvatic habitats in the Loja and Manabí provinces in Ecuador, where they are involved in the transmission of *Trypanosoma cruzi*, the causal agent of Chagas disease. The phylogeny and the phylogeography of these species were studied using the nuclear gene ITS-2 and the mitochondrial gene cytochrome *b*. Phylogeny was explored using Maximum Likelihood (ML) and Bayesian methods. Phylogeography of each species was studied using genetic diversity indices, neutrality and Mantel’s tests. The cytochrome *b* phylogeny showed *P. howardi* and *P. chinai* formed two defined clades, while sequences assigned to *P. rufotuberculatus* were scattered within both clades. The sequences within the two clades diverged from 0% to 3.1%, and between clades from 4.3% to 8.3%, suggesting that *P. chinai* and *P. howardi* are different species, whose divergence might have been assisted by human colonization of the regions. Contrary to the cytochrome *b* data, ITS-2 recovered *P. rufotuberculatus* genotypes within a single clade, with little divergence among haplotypes within the clade (0%-1%). *Panstrongylus howardi* and *P. chinai* were intermixed in a second clade with divergences among haplotypes ranging from 0% to 2.2%. The divergence between the *P.
*rufotuberculatus* and *P. howardi/chinai* clades was substantial (11.7%-18.1%), clearly distinguishing *P. rufotuberculatus* as a distinct species. The discrepancies between the phylogenies of these genes strongly suggest introgression between *P. rufotuberculatus* and the other two species. In addition, the genetic, geographic, morphologic, and biologic characteristics of *P. chinai* and *P. howardi* and their haplotypes suggest that these correspond to two evolutionary units. The phylogeographic results showed no genetic structure among populations of *P. howardi* from the Manabi province. Rather, a recent population expansion event is suggested by neutrality tests. Positive isolation-by-distance among populations located beyond 10 km was found. Because *P. howardi* has been found in sylvatic and domestic/peri-domestic habitats, this information suggests that *P. howardi* re-infests habitats after other species of triatomines have been eliminated by the application of insecticide sprays. In Loja, *P. chinai* showed no intra-specific structure or population expansion. However, a degree of isolation-by-distance was detected, suggesting limited dispersal. *Panstrongylus rufotuberculatus* lacked both genetic structure and isolation-by-distance, suggesting more widespread dispersal between the Loja and Manabi provinces. Such findings suggest that dispersal of the bugs from one province to another might be human assisted. Passive transport of bugs hitchhiking on human clothes and domestic animals has been reported. Discordant demographic inferences between ITS-2 and cytochrome *b* suggests sex-biased dispersal within *Panstrongylus*. 
Approved: _____________________________________________________________

Mario Grijalva

Associate Professor of Biomedical Sciences
To my family
ACKNOWLEDGMENTS

I owe the completion of this project to a number of people who helped me in many ways during the process of my dissertation. I would like to thank:

My academic adviser, Dr. Mario Grijalva, for his advice, and for challenging me throughout all the steps in my academic career.

Dr. Ellen Dotson who gave me the opportunity to work in this project in collaboration with the Centers for Disease Control and Prevention (CDC).

Dr. Shawn Kuchta, for his invaluable technical advice, support, patience, and guidance.

Dr. Sarah Wyatt, for her support, technical help and constructive criticism.

The members of my dissertation committee: Dr. Edwin Rowland and Dr. Kenneth Goodrum for their patience, support, and invaluable help.

Dr. William Romoser, for his guidance through my academic career, unconditional help, and sincere friendship.

Dr. Morgan Vis, for her advice and patience.

Dr. Paula Marcet for her help in sample processing, and technical support in data analysis as well as troubleshooting and constructive criticism at the CDC.

Gena Lawrence MSc. for her laboratory training.

Lcda. Ana Villacis for providing me with the samples, their morphological, and epidemiological information, used for this study.

Lcdo. Cesar Yumiseva for helping in the elaboration of the maps for this study.
Dr. Daryl Lam and Ms. Danette Pratt for their technical support and help in the use of software.

The personnel from the National Chagas Control Program, National Service for Malaria Eradication (SNEM), Ecuadorian Ministry of Health.

Financial support received from UNICEF/UNDP/World Bank/WHO Special Programe for Research and Training in Tropical Diseases (TDR) Research Capability Strengthening group (grant ID number A20785), the Division of Microbiology and Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health (DMID/NIAID/NIH), Children's Heartlink Minneapolis, Plan International Ecuador, Ohio University College of Osteopathic Medicine, PUCE and the Ecuadorian Ministry of Public Health.

The European Community Latin American Network for Research on the Biology and Control of Triatominae (ECLAT) and Chagas Epidemiological Network (EPINET).

Funding agencies did not play any role on the design of the study or in the data analyses and drafting of this dissertation.
# TABLE OF CONTENTS

| Abstract | 3 |
| Acknowledgments | 7 |
| List of Tables | 12 |
| List of Figures | 13 |
| Chapter 1: Introduction | 15 |
| 1.1 Chagas disease biology and epidemiology in Ecuador and Latin America | 15 |
| 1.2 Chagas disease transmission | 16 |
| 1.3 *Trypanosoma cruzi* life cycle | 18 |
| 1.4 Treatment for Chagas disease | 17 |
| 1.5 Prevention | 17 |
| 1.6 Triatomines | 18 |
| 1.7 Triatomines life cycle | 19 |
| 1.8 Behavior | 19 |
| 1.9 *Panstrongylus sp.* found in the Loja and Manabi provinces, Ecuador | 21 |
| 1.10 Triatomines and chagas disease | 22 |
| 1.11 Objectives and hypothesis of the study | 23 |
| 1.11.1 Specific Aim 1: to infer the phylogeny of *P. howardi*, *P. chinai*, and *P. rufotuberculatus*, which are found in Loja and Manabí | 26 |
| 1.11.2 Specific Aim 2: to analyze the genetic diversity of the species of *Panstrongylus* found in Loja and Manabi | 26 |
| 1.11.3 Specific Aim 3: to analyze the geographical dispersal within species of *Panstrongylus* found in Loja and Manabi | 27 |
| Chapter 2: Phylogenetic study of *Panstrongylus sp.* | 28 |
| 2.1 Introduction | 28 |
| 2.2 Materials and Methods | 34 |
| 2.2.1 Sample collection | 34 |
| 2.2.2 DNA extraction | 34 |
| 2.2.3 PCR reactions | 37 |
| 2.2.4 Gel electrophoresis | 38 |
2.2.5 Cloning of ITS-2 fragments ................................................................. 39
2.2.6 Plasmid purification ........................................................................... 40
2.2.7 Plasmid screeening ........................................................................... 41
2.2.8 Plasmid purification for sequencing .................................................. 42
2.2.9 Cyt b cycle sequencing reaction ......................................................... 43
2.2.10 Cyt b cycle sequencing reaction purification ..................................... 44
2.2.11 Cyt b cycle sequencing ..................................................................... 44
2.2.12 ITS-2 plasmid cycle sequencing ....................................................... 45
2.2.13 Data analysis ..................................................................................... 45
2.3 Results .................................................................................................... 46
2.4 Discussion .............................................................................................. 53

Chapter 3: Phylogeographic study of Panstrongylus sp. in the Loja and Manabi provinces, Ecuador ........................................................................................................... 63
3.1 Introduction .............................................................................................. 63
3.2 Materials and Methods ........................................................................... 68
  3.2.1 Genetic diversity indices .................................................................... 68
  3.2.2 Intra-specific genetic structure .......................................................... 69
  3.2.3 Haplotype networks .......................................................................... 69
  3.2.4 Demographic history ........................................................................ 70
  3.2.5 Geographic dispersal ......................................................................... 71
3.3 Results ..................................................................................................... 72
  3.3.1 Panstrongylus howardi ................................................................. 72
  3.3.2 Panstrongylus chinai ........................................................................ 76
  3.3.3 Panstrongylus rufotuberculatus ......................................................... 80
3.4 Discussion ............................................................................................... 84
  3.4.1 Phylogeography of P. howardi ........................................................ 84
  3.4.2 Phylogeography of P. chinai ............................................................. 90
  3.4.3 Phylogeography of P. rufotuberculatus ............................................ 95
  3.4.4 Implications for the transmission of Chagas disease ....................... 98
  3.4.5 Implications for the control of Chagas disease ............................... 99
  3.4.6 The context of this study ................................................................. 101
References.................................................................................................................. 104
Appendices.................................................................................................................. 159
  Appendix 2.1: List of triatomines collected in Manabi and Loja provinces in Ecuador
  and used for the analyses in this study................................................................. 160
  Appendix 2.2: Preparation of plasmid extraction solutions..................................163
  Appendix 2.3: GARLI script used to conduct Cyt b phylogenetic inference using
  Maximum Likelihood method .............................................................................. 164
  Appendix 2.4: GARLI script used to conduct ITS-2 phylogenetic inference using
  Maximum Likelihood method .............................................................................. 166
  Appendix 2.5: MrBayes script used to conduct Cyt b phylogenetic inference using
  Bayesian method ................................................................................................. 168
  Appendix 2.6: MrBayes script used to conduct ITS-2 phylogenetic inference using
  Bayesian method ................................................................................................. 169
  Appendix 2.7: GARLI script used to conduct phylogenetic inference of South
  American ITS-2 samples using Maximum Likelihood method......................... 170
  Appendix 2.8: MrBayes script used to conduct phylogenetic inference of South
  American ITS-2 samples using Bayesian method ............................................... 172
LIST OF TABLES

Table 3.1: Genetic indices among *Panstrongylus sp.* ...................................................123
Table 3.2: $F_{ST}$ calculated from Cyt b among *P. howardi* ..............................................124
Table 3.3: $F_{ST}$ calculated from ITS-2 among *P. howardi* .............................................125
Table 3.4: Neutrality tests and demographic history ..........................................................126
Table 3.5: $F_{ST}$ calculated from Cyt b among *P. chinai* ................................................127
Table 3.6: $F_{ST}$ calculated from ITS-2 among *P. chinai* ................................................128
Table 3.7: $F_{ST}$ calculated from Cyt b among *P. rufotuberculatus* .................................129
Table 3.8: $F_{ST}$ calculated from ITS-2 among *P. rufotuberculatus* .................................130
LIST OF FIGURES

Page

Figure 1.1: Species of *Panstrongylus* included in this study.................................131

Figure 2.1: Map of Ecuador ..........................................................................................132

Figure 2.2: *Panstrongylus* sp. cytochrome b PCR ........................................................133

Figure 2.3: *Panstrongylus* sp. ITS-2 PCR ..................................................................134

Figure 2.4: Plasmid purification ...................................................................................135

Figure 2.5: *EcoRI* plasmid digestion .........................................................................136

Figure 2.6: Bayesian analysis data convergence..........................................................137

Figure 2.7: Cytochrome b phylogram of *Panstrongylus* sp. .........................................138

Figure 2.8: ITS-2 phylogram of *Panstrongylus* sp. ...................................................139

Figure 2.9: Cytochrome b and ITS-2 phylogram of *Panstrongylus* spp. ......................140

Figure 2.10: ITS-2 phylogram of *Panstrongylus* sp. found in South America.............141

Figure 3.1: Map of the Manabí province .....................................................................142

Figure 3.2: Geographic distribution of *P. howardi* Cyt b haplotype network ..............143

Figure 3.3: Geographic distribution of *P. howardi* ITS-2 haplotype network ..........144

Figure 3.4: *Panstrongylus howardi* mismatch distribution ........................................145

Figure 3.5: Isolation-by-distance of *P. howardi* using Cyt b sequences .....................146

Figure 3.6: Isolation-by-distance of *P. howardi* using ITS-2 sequences .....................147

Figure 3.7: Map of the Loja province ..........................................................................148

Figure 3.8: Geographic distribution of *P. chinai* Cyt b haplotype network ...............149

Figure 3.9: Geographic distribution of *P. chinai* ITS-2 haplotype network ...............150
Figure 3.10: *Panstrongylus chinai* mismatch distribution.................................151

Figure 3.11: Isolation-by-distance of *P. chinai* using Cyt b sequences.........................152

Figure 3.12: Isolation-by-distance of *P. chinai* using ITS-2 sequences ......................153

Figure 3.13: Geographic distribution of *P. rufotuberculatus* Cyt b haplotype network
........................................................................................................................................154

Figure 3.14: Geographic distribution of *P. rufotuberculatus* ITS-2 haplotype network
........................................................................................................................................155

Figure 3.15: *Panstrongylus rufotuberculatus* Cyt b mismatch distribution ..................156

Figure 3.16: *Panstrongylus rufotuberculatus* ITS-2 mismatch distribution...............157

Figure 3.17: Isolation-by-distance of *P. rufotuberculatus* using Cyt b sequences ......158

Figure 3.18: Isolation-by-distance of *P. rufotuberculatus* using ITS-2 sequences.......159
CHAPTER 1: INTRODUCTION

1.1 Chagas disease biology and epidemiology in Ecuador and Latin America

Chagas disease, also called American Trypanosomiasis, is distributed in Latin American countries and the Caribbean with approximately 10 million infected people and 50,000 new cases every year (Hotez, 2008; WHO, 2011). Among the infected population, approximately 5.4 million will develop to the chronic stage, and another 900,000 will develop megaesophagus and megacolon (Hotez, 2008). In Ecuador, there are approximately 230,000 infected people and 6.2 million at risk (PAHO, 2006). The intracellular protozoan parasite *Trypanosoma cruzi* is the causative agent of Chagas disease (WHO, 2002). The vectors of *T. cruzi* are blood feeding triatomine insects (WHO, 2002). After feeding on its victim, the triatomine deposits its feces containing infective parasites near the feeding lesion or near the eyes. Later, the victim scratches the area and carries the parasites into the bloodstream through the bite puncture or mucosal surfaces (WHO, 2002). However, blood transfusion, ingestion of contaminated food, or vertical transmission can also result in infection with the parasite (WHO, 2002).

The disease has two phases: the acute and the chronic phase. The acute phase appears right after the infection of the patient with *T. cruzi*, usually during childhood. The symptoms include an initial local reaction to the bite, also called “chagoma.” Among other manifestations of the disease are: fever, lymphadenopathy, splenomegaly, hepatomegaly, and heart failure. These acute signs and symptoms can disappear in 4 to 8 weeks. The signs and symptoms of the chronic phase of Chagas disease generally appear
5 to 20 years after the initial infection. The period of time between the acute and the chronic phases is known as the “indetermined” period, characterized by an apparent absence of symptoms. However, progressive changes in the function of several organs such as heart, nervous system, esophagus, and colon may occur, which mark the appearance of the chronic form. Among these, cardiomyopathy, with a very poor prognosis, has been identified as the most important (WHO, 2002).

1.2 Chagas disease transmission

The most common means of transmission of Chagas disease to humans occurs when infective forms of *T. cruzi*, the trypomastigotes, are egested in the feces of the triatomine after it has taken a blood meal from the host (Miles, 2009; WHO, 2002). The trypomastigotes penetrate the skin through the insect’s bite site, skin abrasions, or mucosa when the person scratches the parasite-laden feces into the wound; that is, transmission of the parasite does not occur through the mouthparts during biting, but only through contamination with the bug’s excrement (Miles, 2009). However, blood transfusion, ingestion of contaminated food, and vertical transmission can also result in human infection with the parasite (WHO, 2002; Miles, 2009).

1.3 *Trypanosoma cruzi* life cycle

Trypomastigotes taken in a blood meal by the insect will develop into epimastigotes. In the midgut within 15 days these forms divide and become metacyclic trypomastigotes. The parasites then locate in the lumen of the bugs’ digestive tract and
malpighian tubules where they are egested after another blood meal (Miles, 2009; CDC, 2010). In humans, after being introduced into the host’s system, infective non-dividing forms of the parasite, trypomastigotes, infect phagocytic cells and also cells from cardiac, gastrointestinal, and adipose tissues, among others (Stuart et al., 2008; Teixeira et al., 2006). After infection, intracellular trypomastigotes develop into amastigote forms, which divide by binary fission and later differentiate into metacyclic trypomastigotes (Stuart et al., 2008). These then burst from the infected cell to infect other tissues.

1.4 Treatment for Chagas disease

Currently, there is no effective treatment for Chagas disease. The drug Nifurtimox is currently used and attacks intracellular forms of the parasite producing a cure rate of about 90%. However, Jackson et al. (2010) found that 97.5% of patients who take this drug present adverse side effects. While most of these effects might be expected and not severe, a small number of patients might present eosinophilia, quincke edema, acute myocarditis and anaphylaxis. The poor tolerance to nifurtimox by the patients strongly influences the low rate of treatment completion, which makes the general use of the drug unpractical (Jackson et al., 2010). The drug Benznidazole has also been used effectively in acute cases. However, its effectiveness in undetermined and chronic cases still remains uncertain (Viotti et al. 2006).
1.5 Prevention

The best approach to stopping Chagas disease transmission is to prevent contact between humans and the insect vector through vector control interventions (Eldridge and Edman, 2004). For this purpose, application of pyrethroid insecticide on the interior of surface walls and roofs of houses has been used. Deltamethrin, a potent type II pyrethroid (Imamura et al., 2006), is one of the insecticides most used (25a.i.mg/m²) to eliminate house-infesting triatomines (Rojas de Arias et al, 2004; Grijalva et al., 2011). The lethal toxicity of this insecticide for insects is attributed to its ability to induce the permanent opening of sodium channels, primarily affecting the insect’s sensory neurons, paralyzing the bug (Vijverberg and vanden Bercken, 1990). However, cases of resistance to the insecticide have been reported (Marquardt, 2005).

1.6 Triatomines

Triatomines (Hemiptera: Reduviidae) are the vectors involved in the transmission of Trypanosoma cruzi, the causative agent of Chagas disease. This parasite was identified in 1909 by Carlos Chagas (Lewinsohn, 2003; WHO, 2002). Chagas disease is a zoonosis, as the parasite is naturally found in opossums and other small wild mammals, which serve as their reservoir hosts (WHO, 2002). Humans become accidentally infected with T. cruzi when they enter the triatomines natural environment, and the rates of infection can increase if wild animals are killed or driven out from their habitats, forcing the triatomines to seek blood meals among domestic animals and humans as opposed to their natural sources (WHO, 1997). Triatomines are characterized by being obligate
hematophagous parasites in males and females, and in all life stages. In addition, these insects have a distinctive shape (flattened) that allows them to hide in crevices (Eldridge and Edman, 2004).

1.7 Triatomines life cycle

The bug’s life cycle starts when eggs are deposited in or near the host resting places, including mammal nests, chicken coops, beds, as well as leaves of trees. Commonly, a female lays about 1000 eggs in its lifetime, doing so in about 1 to 2 eggs per day. After 15 to 30 days, small pale nymphs emerge from eggs and remain hidden for about 2 to 3 days before beginning to seek their first blood meal. There are 5 nymphaal instars, which require at least one blood meal to initiate change to the next developmental stage. Nymphs can ingest up 5 times their own weight in blood, decreasing their proportional demand to 4 times their own weight in a blood meal when adults. They can feed every 4 to 9 days. The complete life cycle of the triatomine from egg to adult takes from 6 to 10 months. However, when the bugs do not have sufficient sources of blood, the cycle can be extended and last up to 2 years (WHO, 1997; Marquardt, 2005).

1.8 Behavior

The bugs bite human hosts at night, taking about 2 mL of blood per person per night. Blood feeding takes 10 to 15 minutes, and can contribute to anemia in victims, in addition to the transmission of T. cruzi (Marquardt, 2005). Every triatomine species is susceptible to T. cruzi infection. However, high rates of natural infection with the parasite
are found in about 70 species (Sainz, et al. 2004). Moreover, triatomines are known for taking blood meals from one another, which contributes to hyperparasitism. In fact, the degree of susceptibility of the species to host the parasite, the time interval between feeding and defecation, and the degree of the bug’s contact with the human host play major roles in determining the vector capacity of triatome species (Eldridge and Edman, 2004).

Three habitats or resting places where triatomines are found have been described: domestic, peri-domestic and sylvatic. Domestic resting places for triatomines are defined as those places where the insects are found dwelling in the closest contact with human hosts (i.e. in the homes). Usually, these places are dark (WHO, 1997; Maquardt, 2005). Among these, places like dark crevices, behind pictures, in furniture, boxes and clothing hanging from walls and pegs, and walls have been observed (WHO, 1997). Peri-domestic resting places are defined as the places where triatomines are found dwelling near the human environment. Usually this term refers to the area around houses where domestic animals might be found. Among the peri-domestic areas, firewood, stored objects, lumber, tiles, stones, bags of food, chicken coops and goat corrals have been observed. Finally, when species of triatomines are found feeding on mammals in the wild, the insect’s habitat is defined as sylvatic (WHO, 1997; Maquardt, 2005). In spite of the defined areas where triatomines can be found, cases of adaptation and habitat change have been reported. For example, sylvatic species might move from forest to houses when their habitat has been affected; meanwhile, other species might appear in both kinds
of habitats: domestic, feeding on farm and domestic animals, and humans; and sylvatic, feeding in the wild (Maquardt, 2005; WHO, 1997).

The availability of warm-blooded sources is the factor that influences the presence of the insects in local habitats. It has been observed that transportation of triatomines mainly depends on humans since the insects frequently hitchhike on people or their belongings (Eldridge and Edman, 2004).

1.9 Panstrongylus sp. found in Loja and Manabi provinces, Ecuador.

This study focuses on the species *P. howardi* (Fig. 1.1A), *P. chinai* (Fig. 1.1B), and *P. rufotuberculatus* (Fig. 1.1C) found in Loja and Manabi provinces, Ecuador. These provinces contain different ecosystems, and the residing bugs present different behaviors and habitat preferences. All three species are found naturally infected with *T. cruzi* (Lent and Wygodzinsky, 1979; Suarez-Dávalos et al., 2010; Grijalva et al., 2011). Only found in wet forests (Patterson et al., 2009) of Manabi province, *P. howardi* has been reported to be one of the two main species found to infest peri-domestic, domestic, and sylvatic habitats in Manabi province (Lent and Wygodzinsky, 1979; Suarez-Davalos et al., 2010; Grijalva et al., 2011). Taxonomically, this species is closely related to *P. chinai*, but the specific phylogenetic relationship between them has not been clearly defined (Patterson et al., 2009). This species has also been misidentified in the field as *T. dimidiata*, due to their similar chromatic pattern (Aguilar et al., 1999; Grijalva et al., 2011).

*Panstrongylus chinai* is found distributed in Venezuela, Ecuador and Peru (Patterson et al., 2009). They have been found in chicken nests and, according to Lent
and Wygodzinsky (1979) and Patterson et al. (2009), occasionally in human habitations. Furthermore, Grijalva et al. (2005) reported the occasional intra-domiciliary presence of *P. chinai*, which later became predominant in domestic habitats in the Loja province, according to Grijalva (unpublished data). This species is generally found in deserts or deserted brush-woods (Patterson et al., 2009). Members of this group are considered by some researchers to be melanic forms of *P. howardi* because of their close phenotypic and genotypic relationship (Lent and Wygodzinsky, 1979; Patterson et al., 2009). Moreover, Grijalva (unpublished data) reports the predominance of *P. chinai* in domestic habitats in Loja province.

*Panstrongylus rufotuberculatus* is found distributed through Brazil, Bolivia, Colombia, Costa Rica, Ecuador, Panama, Peru, and Venezuela (Lent and Wygodzinsky, 1979). Recent studies report their presence in sylvatic and domestic habitats (Grijalva et al., 2005; Patterson et al., 2009) including dry, wet and moist forests, where this species can be found (Patterson et al., 2009).

1.10 Triatomines and Chagas disease

Triatomines are found widely distributed throughout 21 countries in Central and South America (Marcilla et al., 2002). There are approximately 10 million infected people, and another 25 million at risk of infection with Chagas disease (WHO, 2011). In Ecuador, Chagas is classified as an endemic disease (Grijalva et al., 2005). In fact, according to Aguilar et al. (1999), historical reports have documented its presence since
1950. The same source also estimates that approximately 200,000 people were already infected in Ecuador by 1999 (Aguilar et al. 1999).

Triatomine insects have a sylvatic origin, but deforestation and colonization of their environment have changed their natural behavior. Recent conducted studies report the presence of *Rhodnius ecuadoriensis*, *Triatoma dimidiata*, and *T. carrioni* in Southern Ecuador (Grijalva et al., 2005), *Panstrongylus rufotuberculatus*, *P. howardi*, *T. carrioni* and *P. chinai* in El Oro, Manabi, and Loja provinces (Aguilar et al. 1999), and six species of *Panstrongylus*, distributed in the areas of tropical and subtropical climates of Ecuador (Abad-Franch et al., 2001). Thus Triatomines, particularly individuals from the genus *Panstrongylus*, have a wide and diverse geographical distribution in Ecuador.

Chagas disease prevention is focused on the eradication of the vector, but the lack of information about their variety, diversification, and geographic dispersal makes it difficult to establish successful specific interventions. The genera *Rhodnius* and *Triatoma* have been the most genetically studied, and the information provided by those studies has been valuable in improving control programs in South America. However, species in the genus *Panstrongylus* have emerged as potential vectors of *T. cruzi*. Therefore, studying their genealogical history, population genetics, and geographic dynamics in the region is essential to design successful Chagas disease vector control programs. Such information would help researchers to identify species that might be most involved in *T. cruzi* transmission, identify newly established populations of triatomines, and monitor the outcomes of vector control interventions in Loja and Manabi provinces.
1.11 Objectives and hypotheses of the study

In Ecuador, sixteen species of triatomines and their geographical distribution have been described (Abad-Franch et al., 2001). Among them, six species of *Panstrongylus* have been found in the Coastal, Highlands, and Amazon regions of Ecuador (Abad-Franch et al., 2001). Recently, species of the genus *Panstrongylus* have been found in numerous localities of Loja and Manabí provinces, Ecuador (Grijalva et al., 2005; Suarez-Davalos et al. 2010; Grijalva et al., 2011; Grijalva, unpublished). Indeed, high infestation levels have been found in domestic and peri-domestic habitats in Loja and in peri-domestic habitats in Manabí province. Further, *Trypanosoma cruzi* infection has also been observed in *Panstrongylus* spp. In addition, recent studies have documented an increase in domiciliation as well as re-infestation of houses previously treated with insecticides (Grijalva, unpublished).

Because of their high propensity for domiciliation and their tendency to fill the ecological niches left empty after control interventions, species of the genus *Panstrongylus* deserve research attention (Patterson et al., 2009), including consideration of intra- and inter-specific genetic and geographical relationships among Ecuadorian species in this genus. On this basis, the general objective of my research has been to infer the phylogeny of *Panstrongylus*, and study the phylogeography of three *Panstrongylus* species collected in Loja and Manabí. The results will contribute to the design of better triatomine control strategies for this region. In addition, it is hoped that the tools developed for these analyses will serve to evaluate the effectiveness of current and future vector control interventions.
1.11.1 Specific Aim 1: to infer the phylogeny of P. howardi, P. chinai, and P. rufotuberculatus, which are found in Loja and Manabí.

Previous studies conducted in Latin America have inferred the phylogeny of species of *Panstrongylus* using mitochondrial or nuclear DNA (Lyman et al. 1999; Marcilla et al. 2002; Hypsa et al. 2002). In addition, previous studies conducted by Lent and Wygodzinski (1979) using morphological characteristics, and Marcilla et al. (2002) using nuclear DNA (ITS-2), revealed some of the genealogical relationships and the polyphyletic nature of the species in the genus *Panstrongylus*. However, phylogenetic analysis of species within *Panstrongylus* using both mitochondrial and nuclear DNA together to define the genealogy of this genus, including clarification of the relationship between the species *P. howardi* and *P. chinai*, has not been conducted. It has been suggested that *P. chinai* might represent a melanic form of *P. howardi* (Patterson et al., 2009). I have tested the hypothesis that *P. howardi*, *P. chinai*, and *P. rufotuberculatus* found in Loja and Manabí are three well genealogically defined, monophyletic species.

1.11.2 Specific Aim 2: to analyze the genetic diversity of the species of *Panstrongylus* found in Loja and Manabí.

Species in the genus *Panstrongylus* have been found to be natural carriers of *T. cruzi*. They are found dwelling in domiciliary, peri-domiciliary, and sylvatic habitats in Manabí, and in domiciliary and peri-domiciliary habitats in Loja provinces. Vector control programs directed against populations of triatomines in rural areas of these provinces have been implemented. However, re-infestation has been reported after
insecticide spraying (Grijalva et al. 2011). The degree of genetic diversity at the intra-specific levels of *Panstrongylus* might serve as an indicator of the population demographic status as well as the possible effects of vector control programs on the genetic variability of the bug populations. Such information about the intra-specific diversity of *Panstrongylus* sp. can be used in subsequent planning of vector control strategies. The methodology used in this project can also serve as a tool for monitoring the effects of vector control interventions on triatomine populations by measuring the degree of genetic variability among populations of *Panstrongylus* sp. and the possible demographic changes that could have influenced the observed results.

1.11.3 Specific Aim 3: to analyze the geographical dispersal within species of *Panstrongylus* found in Loja and Manabi.

Species of *Panstrongylus* are geographically distributed according to specific life zones. For example, the presence of *P. rufotuberculatus* has been reported in dry, wet, and moist forests; *P. chinai* has been found in deserts, deserted brushwood or steppe and thorny brushwood, and *P. howardi* has been found in wet forests (Patterson et al., 2009). It has been stated that human intervention has helped the passive dispersal and colonization of new areas by triatomines (Noireau et al., 2009). In fact, according to Eldridge and Edman (2004), triatomines have a tendency to “hitchhike” on people or within their belongings. Dispersal of *Panstrongylus* species within Ecuador has not been described. Testing isolation-by-distance within these species will help in assessing whether more distant populations within a species are also more genetically different, that
is the hypothesis that geographic distance is proportional to the genetic distance among populations of each species of *Panstrongylus* found in Ecuador. This information will be useful for estimating the degree of isolation of populations within species of *Panstrongylus* as an indicator of susceptibility to vector control programs. The goal is to identify communities within a province where populations of *Panstrongylus sp.* can be isolated due to their geographic distance from other communities.
2.1 Introduction

The study of the evolution of a species focuses on its environmental, morphological, biological, and genetic information, given that the phenotype of an organism is the result of the interaction between the environment and the genes passed through generations (Templeton, 2006). Phylogeny is the area of systematics that estimates the genealogical history of a species. This estimation can use genetic or character data to establish a hierarchy or relationships among individuals with homologous characteristics (inferred common ancestry). To correctly infer phylogeny from nucleotide data, the DNA sequence data used must have descended directly from a common ancestor (orthologous), where the common ancestry of the sequence can be traced back to a speciation event (Moritz and Hillis, 1996).

Mitochondrial DNA (mtDNA) has been widely used to infer phylogeny and conduct phylogeographic studies in different organisms, including Triatomines (Lyman, 1999; Avise, 2000; Templeton, 2006). The complete mtDNA genome of *Triatoma dimidiata* (Hemiptera: Reduviidae) has been sequenced and described by Dotson and Beard (2001). This genome contains 17,019 bp (base pairs), including thirteen protein coding regions, 22 tRNAs, small and large ribosomal units, and a control region (Dotson and Beard, 2001). Because the genus *Triatoma* is closely related to the genus *Panstrongylus*, is likely that the structure and organization of its mitochondrial DNA does not differ substantially.
Mitochondrial DNA is considered as an appropriate region of the DNA to be analyzed for conducting phylogenetic studies because of its maternal transmission, extensive intra-specific variation, and lack of genetic recombination (Avise, 2000). One of the most frequently used regions for phylogenetic studies within mitochondrial DNA is the Cytochrome b (Cyt b) region. The region is characterized by a bias towards A-T content in insects (Page and Holmes, 1998). In *Triatoma*, the region corresponding to Cyt b (10252–11383 bp) is located between the NADH -6 component and a tRNA (S) unit. In addition, Cyt b has been useful for phylogeographic studies of closely related species and phylogenetic reconstructions at the generic and family levels (Avise, 2000).

The internal transcribed spacer-2 (ITS-2) locus, part of the nuclear ribosomal DNA, was used in this study to complement the information given by Cyt b. There are three nuclear ribosomal DNA genes organized in a cluster that contains two internal transcribed spacers (ITS-1 and ITS-2) located between these genes. This complex, as well as an external transcribed spacer (ETS) unit, is tandemly repeated hundreds to thousands of times in the eukaryotic genome. A non-transcribed spacer that separates the clusters along the chromosome is located between these tandem repeats. In terms of conservation of this DNA, the regions corresponding to the genes are more conserved than the internal transcribed spacers, and the internal transcribed spacers are more conserved than the non-transcribed region (Palumbi, 1996).

The internal transcribed spacer 2 (ITS-2) region has gained particular attention among phylogenetic and population studies in Triatomines (Marcilla et al., 2001; Marcilla et al., 2002; Dorn et al., 2009). In fact, the hyper-variable nature of this non-
coding region of DNA makes it a suitable marker to resolve close phylogenetic relationships and micro-evolutionary events. The A-T rich ITS-2 region is located between the 5.8S and 28S units of the ribosomal DNA (rDNA). In addition, tandem repeats of mono-, di-, or tri- nucleotides can be observed within this region. Usually, these tandem repeats, also called microsatellites, influence the length of the ITS-2 fragment amplified from different samples. These microsatellites have been previously used to conduct population studies (Marcet et al., 2006).

Triatomine insects belong in the order Hemiptera, the “true bugs”, family Reduviidae, and subfamily Triatominae (Marquardt, 2005). This subfamily has 6 tribes (Alberproseiini, Bolboderini, Cavernicoloni, Linshcostenini, Rhodniini and Triatomini) divided into 18 genera: Alberprosenia, Belminus, Bolbodera, Microtriatoma, Parabelminus, Cavernicolai, Linshcosteus, Psammolestes, Rhodnius, Dipetalogaster, Eratyrs, Hermanlentia, Meccus, Mepraia, Nesotriatoma, Paratriatoma, Panstrongylus, and Triatoma (Costa et al., 2008; Patterson et al. 2009; Noireau et al. 2009). There are 140 recognized species within the Triatominae (Noireau et al. 2009). The genus Panstrongylus, the third largest genus within the Triatominae subfamily, is composed of 13 species distributed within Central America and South America (Patterson et al. 2009). The diversity of this genus is reflected by the diversity of ecological preferences shown by its different species (Lent and Wygodzinsky, 1979).

Previous studies conducted in Latin America have inferred the phylogeny of species of Panstrongylus using mitochondrial and nuclear DNA (Monteiro et al. 1999; Marcilla et al. 2002; Hypsa et al. 2002). In addition, previous studies conducted by Lent
and Wygodzinsky (1979) using morphological characteristics, and Marcilla et al. (2002) using nuclear DNA (ITS-2), revealed the inter-specific genealogical relationships and a possible polyphyletic nature of species in *Panstrongylus* (the species in the genus evolved from different ancestors). However, a study involving the phylogenetic inference of species of *Panstrongylus* using both mitochondrial and nuclear DNA together to infer the genealogy of *Panstrongylus* has not been conducted. In this case, the relationships between the closely related *P. howardi* and *P. chinai*, currently defined as two different species, is of special interest since it has been suggested that *P. chinai* might represent a melanic form of *P. howardi* (Patterson et al., 2009). Therefore, my study aimed to infer the phylogeny of *P. howardi*, *P. chinai*, and *P. rufotuberculatus* in the Loja and Manabí provinces.

Different types of morphological and DNA data and analytic approaches have been used to conduct phylogenetic inferences, depending on the questions being posed by the researcher. Genomic DNA sequence data are considered a representation of an organism genome, which increases the accuracy of phylogenetic studies (Cumming et al., 1995); although, one nuclear gene does not describe the entire genome because different genes could experience different histories (Knowles and Kubatko, 2010). Therefore, these data may be more suitable for conducting studies at the level of most evolutionary events (Hartl, 2000). Among the available methods to conduct phylogenetic analyses, maximum parsimony, maximum likelihood, and Bayesian analyses are considered the most robust methods. Maximum parsimony (MP) was one of the first approaches to infer phylogeny when only morphological data was used. These methods were adapted to the
development of technology that allowed the collection of data from genetic material and development of software able to handle large amounts of data. Maximum parsimony infers phylogeny using the optimality criterion of finding a tree with the least amount of character change. In the case of using DNA sequence data, MP conducts inference on the basis of nucleotide differences among sequences. However, according to Huelsenbeck and Lander (2003) and Felsenstein (1978), parsimony becomes inconsistent when evolutionary rates vary greatly along branches due to “long branch attraction.”

Maximum likelihood (ML) methods of parameter estimation (Nei, 1996) use models of evolutionary change to infer phylogenies from a set of hypotheses, which include different tree structures, branch lengths, the parameters of the model of sequence evolution, etc. Maximum likelihood gives the smallest variance of a parameter estimate when the sample size is large (Nei, 1996). Thus, testing a hypothesis in ML is testing the probability of a dataset given a stochastic evolutionary model and branch lengths (Wheeler, 2010). According to Swofford et al. (1996), this method is less affected by sampling error than other methods. According to Gaut and Lewis (1995), ML performs as well or better than most parsimony or distance-based methods with statistical consistency and minimum variance, even when the ML assumptions are violated. However, this method requires more computational burden compared to other methods (Hasegawa and Kishino, 1994). Bootstrap resampling is used to evaluate the accuracy of the ML, as well as the MP. This process resamples the original dataset by conducting random character replacements (process repeated hundreds of times) and reconstruct trees for later observe
the frequency of a particular branch tree with the highest-likelihood among the other
trees, after a consensus tree has been reconstructed. (Hasegawa and Kishino, 1994).

Bayesian inference of phylogeny is similar to ML in that it uses a likelihood
function and an explicit model of nucleotide substitution, but differs in that Bayesian
methods include prior knowledge in the hypothesis testing in addition to the data
(Archibald et al. 2003) to calculate a posterior probability distribution of trees
(Templeton 2006; Lemey et al., 2009). The prior probability is defined as the probability
of different parameters given previous knowledge; posterior probability is defined as the
probability of parameters considering a prior distribution, model, and data (Archibald et
al., 2003). The value of the posterior probability establishes the probability that a clade is
“true” given the prior model and the data (Archibald et al., 2003), and is obtained by
combining the prior probability and the likelihood for each tree (Huelsenbeck et al.,
2001). The Bayesian method most commonly used is the one implemented in the
software MrBayes (Huelsenbeck and Ronquist, 2001, Ronquist and Huelsenbeck, 2003;
Altekar et al., 2004).

The use of an evolutionary model of nucleotide substitution by ML and Bayesian
analyses confer these methods the advantage of calculating a “corrected” genetic distance
among taxa. Because the number of substitutions between two DNA sequences does not
keep a directly proportional relationship with time since divergence, uncorrected
distances underestimate the evolutionary change (Page and Holmes, 1998). For this
reason, it is important to define the best-fit nucleotide substitution model for the dataset
under study.
2.2 Materials and Methods

2.2.1 Sample collection

A total of 71 samples of *Panstrongylus sp.* were collected, including 30 from Loja and 41 from Manabí (Figure 1.1). The bugs were captured by intra- and peri-domiciliary searches conducted in each household of the studied communities in Loja and Manabí using the one person-hour method (Gurtler et al., 2001). Samples of *P. howardi*, *P. chinai*, and *P. rufotuberculatus* were collected from 14 rural communities in Manabí (La Cienega, El Bejuco, Zapallo, La Encantada, Maconta Abajo, San Gabriel, San Francisco, Zapallo, Santa Rosa de las Palmas, San Ramon, Estero Seco, Pacoche, Punta Larga, and La Humedad), and from 15 communities in Loja (Algarobillo, Ashimingo, Bramadero, Coamine, Chaquizca, Chirimoyos, La Extensa, Guara, Jaguay, Lucarqui, Naranjo Dulce, Pitayo, San Francisco, El Sauce, and Santa Ester). The samples were identified to the species level by entomologists at the Centro de Investigación de Enfermedades Infecciosas (CIEI) at Quito, Ecuador, according to the morphological characteristics described by Carcavallo et al. (1998) and Lent and Wygodzinsky (1979). Samples of the insect’s legs were preserved in cryovials containing technical grade ethanol. Appendix 2.1 shows samples coding and habitat information.

2.2.2 DNA extraction

Genomic DNA was extracted from each sample, using two legs from each bug. DNeasy blood and tissue kit (Qiagen, Valencia, CA; Promega, Madison, WI; Cat # A1120), and Promega Wizard Genomic DNA Purification kit (Promega, Madison, WI)
were used according to the manufacturer’s recommendations. Briefly, using the Qiagen kit, legs were taken out of the ethanol and left to dry on the bench top overnight. Dried legs were placed in 1.5 mL micro-centrifuge tubes containing 180 μL of sterile PBS. The legs were then ground using a disposable plastic grinder. After the tissue was reduced, 20 μL of proteinase K was added to each sample. Then, samples were incubated at 55 °C overnight. The next day, samples were mixed thoroughly with a vortex for 15 seconds, and 200 μL of Alkaline Lysis buffer was added. After mixing the buffer with a vortex, samples were incubated for 10 minutes at 70°C. After incubation, 200 μL of 95% ethanol were added before a last step of mixing with a vortex. Then, samples were transferred into a DNeasy Mini spin column. The columns were centrifuged for 1 minute at 9,000 rpm. The flow-through was discarded and the collection tube (provided by the manufacturer) was replaced by a new one. Then, the column was washed by adding 500 μL of washing buffer (AW1) and centrifuging the column at 9,000 rpm for 1 minute. Once more, the flow-through was discarded and the collection tube (provided by the manufacturer) was replaced by a new one. A second washing step was conducted using the washing buffer (AW2), provided by the manufacturer. Centrifugation of the column was carried out at 14,000 rpm for 3 minutes. Flow-through and collection tubes were discarded and the column was placed into a new sterile 1.5 mL micro-centrifuge tube. Samples were then eluted by adding 200 μL of Elution Buffer (AE) directly onto the column’s membrane. The column was incubated at room temperature for 1 minute. Finally, the micro-centrifuge tubes containing the columns were centrifuged at 9,000 rpm
for one minute to elute the extracted DNA. This process was conducted at the Tropical Disease Institute Laboratory at Ohio University.

Using the Promega Wizard Genomic DNA Purification kit, two ethanol-preserved legs were left to dry overnight on the laboratory bench top. Dried legs were placed into a Kontes grinding tube all together with a Kontes pellet pestle. The tubes were placed into a Grinder for 15 seconds to pulverize the sample. Then 300 μL of nucleic acid lysis solution was added. The tubes were mixed with a vortex and incubated for 30 minutes at 65ºC. After incubations, tubes were briefly centrifuged to remove liquid from the sides of the tube. Then, 1.5 μL of RNase solution was added and the tube was inverted three times for mixing. Mixed tubes were then incubated at 37ºC during 30 minutes and left cool to room temperature for 5 minutes. After incubations, 100 μL of protein precipitation solution were added and the tubes were vigorously mixed with vortex for 20 seconds. Samples were put on ice for 5 minutes and centrifuged for 20 minutes at 14,000 rpm at 4ºC. The supernatant was carefully transferred to a new sterile 1.5 mL micro-centrifuge tube containing 300 μL of 100% isopropanol. The contents of the tubes were mixed by inverting them carefully 5 times. After mixing, the tubes were centrifuged at 14,000 rpm for 30 minutes at 4ºC. After centrifugation, the supernatant was discarded and 300 μL of 70% ethanol were added. The tubes were again inverted 5 times to mix the contents. Then, tubes were centrifuged again at 14,000 rpm for 10 minutes at 4ºC. The supernatant was discarded, the tube was inverted on a clean piece of absorbent paper, and the pellet was air dried for 20 minutes. After the pellet was dried, 50 μL of Tris EDTA was added and the samples were incubated for 1 hour at 65ºC with agitation of the pellet every 15
minutes. Extracted DNA samples were then stored at 4°C. This process was conducted at the Entomology Branch, Division of Parasitic Diseases and Malaria, Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia.

2.2.3 PCR reactions

The products of the DNA extraction were used in PCR amplification of specific loci. DNA fragments of the Cytochrome b (Cyt b) region were amplified. For this purpose, the primers 5’-GGA CGW GGW ATT TAT TAT GGA TC-3’ (Cytb7432 forward) and 5’-GCW CCA ATT CAR GTT ART AA-3’ (Cytb7433 reverse) were used (Dorn et al. 2009; Monteiro et al., 2003). For these fragments the same concentrations and components were used in all of the PCR amplification reactions as follows: the total volume of 50 μL of the mix included a final concentration of 0.41 pM of each primer (forward and reverse), 0.025 U of Go Taq polymerase (Promega, Madison, WI), DNTPs, MgCl₂, 2 μL of template DNA, and water. The reactions were set up in a thermal cycler (BIO RAD Laboratories, Inc.) with an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 94°C for 30 sec, 47°C for 30 sec, and 72°C for 1 min, (Monteiro et al. 2004). The PCR reaction was followed by 1 cycle of 72°C for 10 min and a final temperature 4°C. This process was conducted at the CDC.

Forward (F58T 5’-CTA AGC GGT GGA TCA CTC GG-3’) and reverse (R28T 5’-GCA CTA TCA AGC AAC ACG ACT C-3’) primers were used (Marcilla et al. 2001; Dorn et al. 2009) for the amplification of the ITS-2 fragments. The PCR mix for the amplification of the ITS-2 fragments had the same concentrations and components for all
the reactions as indicated: the total volume of 12 μL of the mix included a final concentration of 2.5 U HotStarTaq DNA polymerase, 1.5 mM MgCl₂, 200 μM of each dNTP (Qiagen, Valencia, CA), 0.5 μM each primer, 1μg/50 μL DNA template, and water. The conditions for the PCR reaction were set up in a GeneAmp PCR System 9700 thermal cycler at 95°C for 15 min, followed by 35 cycles of 94°C for 30 sec, 57°C for 30 sec, and 72°C for 1 min. Finally, the PCR reaction was finished with one cycle of 72°C for 10 min and a final temperature of 4°C to stabilize the amplified fragments. The presence of amplified PCR products was confirmed for all the samples by gel electrophoresis in 1% agarose. This process was conducted at the Tropical Disease Institute Laboratory (TDI) of Ohio University (OU).

2.2.4. Gel electrophoresis

Agarose gel was prepared to a concentration of 1% in TAE (Tris-acetate EDTA) buffer and diluted 10 μL of Gel Red (Phenix Research Products, NC) in 100 mL of gel for DNA visualization. Samples were mixed with DNA loading buffer (Promega, Madison WI) at a concentration of 5:1 and loaded onto the gel. Electrical current was set to 100 V for 40 min. Gels were observed using a UV transluminator (Biorad, Hercules, CA), and the presence of PCR products was confirmed for all the samples after observing amplified bands corresponding to 685 bp. This process was conducted at the CDC. For the visualization of the ITS-2 fragment, agarose gel was prepared to a concentration of 1% in TAE buffer. Samples were mixed with DNA loading buffer (Promega, Madison WI) at a concentration of 5:1 and loaded onto the agarose gel. Electrical energy was set to
100 V for 40 min. Gels were observed using a UV transluminator (Biorad, Hercules, CA). The presence of PCR products was confirmed for all the samples after observing multiple amplified bands of variable length. This process was conducted at the Tropical Disease Institute Laboratory of Ohio University.

2.2.5. Cloning of ITS-2 fragments

The cloning process was conducted by using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA). Briefly, a ligation reaction was conducted by mixing, in the following order, 3 μL of PCR product, 1 μL of water, 1 μL of salt solution (1.2 M NaCl and 0.06 M MgCl₂), and 1 μL of TOPO vector (10ng plasmid DNA), to complete a final volume of 6 μL. The mix was prepared on ice, following indications of the manufacturer. After the reaction was prepared, it was left to incubate for 15 minutes at room temperature. While the ligation reaction was taking place, vials containing OneShot chemically competent *E. coli* was left to thaw in ice. After incubation, 2 μL of ligation reaction were transferred into one vial of chemically competent cells. The mixture was gently mixed and put back on ice for 20 minutes. After incubation, cells were transformed by heat shock by placing them in a 42°C water bath for 90 seconds. After heat shock, cells were immediately transferred back into ice and left during 5 minutes. After ice incubation, 250 μL of an enriched bacterial culture medium (S.O.C.) were added to the cells. Then, tubes were transferred to a 37°C thermal shaker, where the samples were mixed at 200 rpm at 37°C for one hour. While cells were incubating, 0.8 mg X-Gal was spread over the surface of LB plates containing 50 μg/L of Kanamicin.
The plates were warmed up in a 37°C VWR incubator (West Chester, PA) before being inoculated with bacteria. After bacterial cells were grown for one hour, these were spread onto the surface of the pre-warmed LB plates using a metallic flame-sterilized cell spreader. Each sample was inoculated onto two plates at two different amounts of bacterial broth: 20μL and 40μL. Finally, the plates were incubated in a VWR incubator (West Chester, PA) for 18 to 24 hours at 37°C. This process was conducted at OU.

2.2.6. Plasmid Purification

After incubation, colonies of transformed bacteria were identified by the white color, indicating successful transformation. Therefore, white colonies were selected over blue colonies for plasmid purification. The alkaline lysis plasmid extraction method was used to screen bacteria containing the desired PCR product. For this purpose, white colonies were picked using a sterile wooden toothpick. Then the toothpick was placed in a plastic cell culture tube containing 5mL of sterile Luria Broth, Miller (Fisher, Waltham, MA) and left inside. The tube was incubated in a thermal-shaker at 37°C for 18-24 hours. The next day, a total of 3 mL of LB broth containing bacteria were centrifuged using a 5415c Centrifuge (Eppendorf North America, Hauppauge, NY) in a 1.5 mL sterile micro-centrifuge tube at 7,000 rpm for 5 minutes. After draining the LB broth from the samples, 100 uL of Solution 1 (See appendix 2.2 for solution components) were added, and the bottoms of the tubes were run along a plastic tube rack to re-suspend the pellet. Following, 200 μL of Solution 2 (See appendix 2.2) were added to each sample and the tubes were mixed by inverting them 10 times. After 5 minutes of incubation in ice, 150
μL of cold Solution 3 (See appendix 2.2 for solution components) were added and gentle mixing was conducted right after the addition of the solution to each sample. After a second incubation for 5 minutes in ice, samples were centrifuged for 5 minutes at 14,000 rpm. The supernatant was transferred to a new 1.5 mL micro-centrifuge tube by pipeting, and 250 μL of phenol was added. After mixing the samples with a vortex, they were centrifuged for 10 minutes at 10,000 rpm. Next, isopropanol was added to fill the total volume of each 1.5 mL micro-centrifuge tube, and the tubes were kept for 20 minutes at -20°C. Later, samples were centrifuged for 10 minutes at 10,000 rpm and the supernatant was discarded. The samples were then washed by adding 200 μL of 70% ethanol. Tubes were centrifuged for 5 minutes at 10,000 rpm, and the supernatant was discarded by pipeting. Samples were then let dry on the bench top during 1 hour, and the pellet was re-suspended in 30 μL of water. Finally, the pellet was allowed to dissolve for 15 minutes, and samples were stored at -20°C until used. This process was conducted at OU.

2.2.7. Plasmid screening

In order to determine the insertion of the desired ITS-2 fragment in a plasmid, plasmids were digested with EcoR1 restriction enzyme (New England BioLab, Ipswich, MA), which has specific targets in the vector before and after the inserted fragment. The 20 μL digestion mix was prepared in the same amount and components for all the samples as follows: 14 μL of water, 1 μL EcoR1 digestion enzyme, 2 μL EcoR1 Buffer, and 3 μL purified plasmid. The mix was then incubated in a VWR (West Chester, PA) incubator for 1 hour at 37°C. The cut fragments were observed and evaluated by running
them in a 1% agarose gel. This process was conducted at the TDI Laboratory of Ohio University.

2.2.8. Plasmid purification for sequencing

After observation in an agarose gel, bacterial colonies holding the desired ITS-2 fragment were again purified using the Qiaprep Miniprep kit (Qiagen, Valencia CA), according to the instructions of the manufacturer. Briefly, a total of 2 mL of LB broth containing bacteria from selected colonies were centrifuged in a 1.5 mL sterile micro-centrifuge tube at 9,000 rpm during 3 minutes. The supernatant was discarded. The bacteria were re-suspended by adding 250 μL of P1 buffer, and mixed by running the bottoms of the tubes along a plastic tube rack. Following, 250 μL of buffer P2 were added and the tubes were inverted 5 times for mixing. Later, 350 μL of N3 buffer were added and mixing was conducted by inverting the tubes 5 times as soon as N3 was dispensed in each sample. Tubes were centrifuged for 10 minutes at 13,000 rpm in a Centrifuge 5415c (Eppendorf North America, Hauppauge, NY) and the supernatant was poured to a Qiaprep spin column. The column was then centrifuged for 1 minute at 13,000 rpm, and the flow was discarded. The spin column was washed with 500 μL of buffer PB, and centrifuged 1 minute at 13,000 rpm. The residual flow was discarded. A second washing step was conducted by adding 750 μL of buffer PE to the column, and centrifuging it 1 minute at 13,000 rpm. The residual flow was discarded and the columns were centrifuged again an additional minute at 13,000 rpm to remove residual buffer. Finally, each column was transferred to a sterile 1.5 mL micro-centrifuge tube. Fifty
micro liters of EB buffer were added directly to the membranes of the columns to eluate the DNA. The samples were left to incubate for one minute at room temperature before a final centrifugation at 13,000 rpm for 1 minute. Collected purified plasmids were stored at -20°C until used. This process was conducted at OU.

2.2.9. Cytochrome b cycle sequencing reaction

The sequences of the Cytochrome b fragments were obtained directly from PCR product cycle sequencing. For the cycle sequencing reaction mix, the forward and reverse primers used for the PCR amplification of the Cyt b fragments from genomic DNA were used in separated reactions. The reaction mix for the PCR cycle sequencing used dye terminator chemistry (Big dyes™, Applied Biosystems, Foster City, CA). The 10 μL mix for the cycle sequencing reactions had the same amounts and components for all the reactions as follows: 0.5 μL of Terminator Ready Reaction Mix v1.1, 2 μL of ABI 5x sequencing buffer, 1μL primer (5pmol), 1μL DNA template (purified PCR product), and 5.5 μL ultrapure water. The thermal cycler conditions were set to 25 cycles: 96 ºC for 10 sec, 50 ºC for 5 min, and 60 ºC for 4 min. The reactions were carried out in a Biorad (Hercules, CA) thermal cycler, at the Entomology Branch, Division of Parasitic Diseases and Malaria, Centers for Disease Control and Prevention (CDC).

2.2.10. Cyt b cycle sequencing reaction purification

After the cycle sequencing reaction took place, reactions were purified using Centri-sep columns (Princeton Separations, Adelphia, NJ), as indicated by the
manufacturer. Briefly, columns were gently tapped to ensure settling of the solid gel on the bottom of the tube. Following, 800 μL of highly pure water were added, and each tube was vigorously mixed with vortex. Hydrated columns were left for 30 minutes room temperature. After hydration was finished, bubbles inside the column were removed by inverting the tube and sharply tapping the tube. The interstitial fluid inside the column was then removed by gravity after taking the upper and bottom caps off the tube. The liquid was completely removed by centrifugation at 3,000 rpm for 2 minutes. Then, cycle sequencing reaction material was added to different columns corresponding to each sample by directly dispensing 20 μL of the reaction mixture onto the center of the gel bed inside the column. Individual columns were placed in collection tubes and centrifuged at 3,000 rpm for 2 minutes. The purification was conducted at the CDC.

2.2.11. Cyt b cycle sequencing

After purification, cycle sequencing reactions were collected into a 96 well plate. The plate was incubated in a thermal block for 10 minutes at 95 °C before running the samples in the sequencer. The cycle sequencing reactions were analyzed in an ABI 3100 DNA sequencer ( Applied Biosystems, Foster City, CA) at the CDC.

2.2.12 ITS-2 plasmid cycle sequencing

ITS-2 purified plasmids were sent to the Ohio University Genomics Facility to be sequenced using the M13F (5’-GTA AAA CGA CGG CCA G-3’) and M13R (5’-CAG GAA ACA GCT ATG AC-3’) primers (Invitrogen, Carlsbad, CA).
2.2.13. Data Analysis

2.2.13.1 Sequence alignment

Sequenced forward and reverse fragments were assembled and examined using the DNASTAR Lasergene software (DNASTAR Inc., Madison, WI), Invitrogen Vector NTI software (Invitrogen Carlsbad, CA), Applied Biosystems Sequence Scanner software (Applied Biosystems, Foster City, CA), and Geneious (Biomatters, Auckland, New Zealand). Sequenced fragments were compared to existing sequences in Genbank using BLAST (Basic Local Alignment Search Tool) available at: http://blast.ncbi.nlm.nih.gov/Blast.cgi. Sequences were then manually aligned using the software Bioedit (Hall, 1999). The software Geneious was later used to refine the alignments under the algorithm of MUSCLE (Edgar, 2004). This alignment was set to 8 iterations.

2.2.13.2 Phylogenetic Analysis

Phylogeny was inferred using alignments corresponding to the Cyt b and ITS-2 DNA fragments. Alignments included sequences from the sampled communities as well as ITS-2 fragments published in Genbank: *Panstrongylus chinai* internal transcribed spacer 2 (ITS2), (Ecuador:Instituto de Higiene) Genbank: AJ306547.1; *Panstrongylus geniculatus* internal transcribed spacer 2 (its2), (Brazil:Para) Genbank: AJ306544.1; *Panstrongylus geniculatus* internal transcribed spacer 2 (ITS2), (Ecuador:Orellana, Yasuni) Genbank: AJ306543.1; *Panstrongylus herreri* internal transcribed spacer 2 (ITS2), (Peru) Genbank: AJ306551.1; *Panstrongylus herreri* internal transcribed spacer 2
(ITS2), (Ecuador: Orellana) Genbank: AJ306550.1; *Panstrongylus lignarius* internal transcribed spacer 2 (ITS2), (Brazil: Para, Santa Barbara) Genbank: AJ306549.1; *Panstrongylus lignarius* internal transcribed spacer 2 (ITS2), (Ecuador: Sucumbios) Genbank: AJ306548.1; *Panstrongylus megistus* internal transcribed spacer 2 (ITS2), (Brazil: minas Gerais, Pampulha) Genbank: AJ306542.1; *Panstrongylus megistus* internal transcribed spacer 2 (ITS2), (Brazil: Belo Horizonte) GENBANK: AJ286886.2; *Panstrongylus rufotuberculatus* internal transcribed spacer 2 (ITS2), (Colombia: Santander, Municipio del Carmen) Genbank: AJ306546.1; *Panstrongylus rufotuberculatus* internal transcribed spacer 2 (its2), (Ecuador: El Oro, Guayacon) Genbank: aj306545.1), published by Marcilla et al. (2002). The analysis also used sequences of *Triatoma sp.* as out-groups to analyze both cyt b and its-2 alignments (*Triatoma dimidiata* isolate MXYUME03 cytochrome b (CYT B) gene, partial CDS; mitochondrial Genbank: FJ197159.1) published by Dorn et al. (2009), (*Triatoma infestans* internal transcribed spacer 2 (ITS2) GENBANK: AJ289876.1) by Marcilla et al. (2001), and (*Rhodnius prolixus* isolate prVE4b cytochrome b (cytb) gene, partial cds; mitochondrial) by Monteiro et al. (2003).

The selection of the best-fit substitution model for the data was conducted using the software jModeltest ver. 0.1.1 (Posada, 2008). This program was set to calculate the maximum likelihood of the data under 11 nucleotide substitution schemes. In addition, the analysis also tested equal and unequal base frequencies, a proportion of invariant sites, and rate of variation among sites. A total of 88 different substitution models were tested (Posada, 2008). Based on the Akaike Information Criterion (AIC), which provides
a measure of fit between model and data (Sullivan and Joyce, 2005), the model selected for the Cyt b dataset was HKY + Γ and for the ITS-2 dataset was HKY.

Maximum Likelihood (ML) and Bayesian analyses were conducted using GARLI ver. 0.96b8 (Zwickl, 2006), and MrBayes ver. 3.1.2 (Huelsenbeck and Ronquist, 2001), respectively. Maximum likelihood analysis using GARLI was set up for Cyt b (Appendix 2.3) and ITS-2 (Appendix 2.4) using the nucleotide substitution models indicated by jModeltest for each dataset. For a preliminary search of the random tree where the search for the most likely tree will start (seed), the first run was set up without bootstrapping. Later, the seed was used for a second run. Bootstrapping in GARLI was set to 1,000 repetitions. A consensus tree was obtained using the 50% majority-rule as implemented in the software Mesquite ver. 2.74 (Madison and Madison, 2010). Bayesian analyses of the Cyt b (Appendix 2.5) and ITS-2 (Appendix 2.6) data were set in MrBayes according to the same models indicated by jModeltest for the two data sets. This method used 2 runs, 3 heated and 1 cold chains of 20,000,000 generations of trees. Trees were sampled every 10,000 generations. Moreover, the first 1,574 trees from the Cyt b data set and 732 from the ITS-2 dataset sampled trees by the Markov Chain Monte Carlo Metropolis Coupling (MCMC) were excluded from the analysis by “burnin”. The exclusion depended on the convergence of the run determined by the value of the standard deviation of the split frequencies. Sample generations were collected until convergence of the stationary distribution of the two runs was reached. Standard deviation of split frequencies < 0.01 was considered as indicative that the two samples have converged
(Lemey et al. 2009). Consensus trees were visualized using the software FigTree 1.6.6 ver. 1.3.1 (Rambaut, 2009).

In addition, a concatenated phylogenetic tree was reconstructed using both Cyt b and ITS-2 alignments. The program MrBayes was used to set up a partitioned run where each DNA fragment was tested using its own evolutionary model, resulting in a single consensus tree from the concatenated results obtained from both markers.

Twelve representative samples of ITS-2 from Loja and Manabi provinces were compared to other *Panstrongylus* sp. found in South America. The alignment was tested for a best-fit model of nucleotide substitution using jModelTest. Using the selected model, phylogeny was inferred using ML (Appendices 2.7) and Bayesian (Appendix 2.8) methods implemented in the already mentioned software.

Finally, genetic distances of the Cyt b and ITS-2 sequences within and between clades in the different trees were calculated to estimate sequence divergence within and between clades. Previously estimated nucleotide substitution models were used to conduct these calculations for each dataset using PAUP* ver. 4.0B. Gamma distribution was included in the calculation of distances for the Cyt b dataset. Dissimilarity between sequences represents the branch length of two taxa in a phylogenetic tree (Swofford et al., 1996).

2.3 Results

Seventy one samples of *Panstrongylus* sp. were sequenced for the Cytochrome b (Cyt b) locus. After PCR amplification, bands of approximately 700 bp were visualized
The Cyt b region was identified using the Basic Local Alignment Search Tool (BLAST) available online at: http://blast.ncbi.nlm.nih.gov/Blast.cgi. Seventy one ITS-2 fragments were amplified, and DNA bands of approximately 800 - 1000 bp were visualized (Fig 2.3). Given the presence of multiple bands, these were inserted into a plasmid, selected (Fig 2.4), and sequenced to obtain fragments of about 700 bp long. Fragments containing the ITS-2 region were compared to fragments published by Marcilla et al. (2002) for identification, using BLAST. Cytochrome b sequences were aligned and truncated to a full length of 641 bp (GenBank Accession Numbers: JX400931 – JX401001). ITS-2 sequences were aligned and truncated to a length of 499 bp (GenBank Accession Numbers: JX400860 – JX400930).

Both Cyt b and ITS-2 alignments were separately tested for the best nucleotide substitution model-fit using the program jModeltest. According to the AIC (Akaike Information Criterion) implemented in the program, the best model-fit for the cytochrome b dataset was HKY+Γ. The Hasenagua Kishino Yano (HKY) model takes into account different rates of change for transitions and transversions; it also allows for unequal equilibrium base frequencies. The gamma (Γ) continuous distribution is used to model among-site rate heterogeneity (Swofford et al., 1996; Page and Holmes, 1998). The nucleotide frequencies were calculated as FreqA = 0.3141, FreqC = 0.2342, FreqG = 0.1150 and FreqT = 0.3367. The transition – transversion ratio was 4.6664, and the shape of the gamma distribution was set at α= 0.1240. The best model selected for the ITS-2 dataset, according to the AIC, was HKY. The nucleotide frequencies were calculated as
FreqA = 0.3262, FreqC = 0.1065, FreqG = 0.1366, and FreqT = 0.4307. The transition-transversion ratio was 1.1119.

Initially, Cyt b and ITS-2 data were separately analyzed using Bayesian and Maximum Likelihood methods. Both analyses supported the same branch structure. However, statistical support varied with each method. Cyt b and ITS-2 datasets were run using MrBayes up to 20,000,000 generations, and a fraction of trees, depending on the generation where data reached convergence, indicated by the value of the standard deviation of the split (Figures 2.6A, 2.6B, and 2.6C), was discarded as “burnin”. Concatenated data were run for up to 80,000,000 generations (Figure 2.6D).

Two sister clades were recovered from the analysis of the three species of *Panstrongylus* found in Loja and Manabí provinces using the Cyt b dataset. *Panstrongylus howardi* and *P. chinai* each formed a separate clade. However, *P. rufotuberculatus* was scattered within both clades. Structure within these clades was, for the most part, recovered as unresolved polytomies. Under the ML analysis, these polytomies were also unresolved, indicating the high similarity among *P. howardi*, as well as among *P. chinai*, samples. The topology of both clades had strong statistical support in terms of posterior probability and likelihood (Fig 2.7). The divergence calculated within the *P. howardi/rufotuberculatus* clade ranged between 0% and 2.0%; and the divergence calculated within the *P. chinai/rufotuberculatus* clade ranged from 0% to 3.1%. Finally, the divergence calculated between the two clades ranged from 4.3% to 8.3%.
Using the ITS-2 dataset, two well-defined and statistically strongly supported clades (posterior probabilities = 1) were reconstructed. One clade was composed of members of *P. rufotuberculatus* except for one individual identified as *P. chinai*. This clade included two lineages of weak statistical support, each including members of both Loja and Manabí provinces. The other clade included members of both *P. howardi* and *P. chinai*. Four lineages were recovered; however, none of them was strongly statistically supported (Fig 2.8). The divergence calculated within the *P. howardi/chinai* clade ranged from 0% - 2.2%; and the divergence calculated within the *P. rufotuberculatus* clade ranged between 0% and 2.2%. Finally, the divergence calculated between the two clades ranged from 11.7% - 18.1%.

To improve resolution, an analysis combining both datasets was conducted using MrBayes. Data were partitioned taking into account the nucleotide substitution models previously selected by jModelTest. The resulting tree recovered a monophyletic clade including *P. rufotuberculatus, P. chinai* and *P. howardi*. The basal clade recovered *P. rufotuberculatus* from Loja and Manabí provinces. *Panstrongylus chinai* and *P. howardi* were recovered in two separate clades, which also defined the geographical locations where the samples were collected. Two samples of *P. rufotuberculatus* (*P. rufotuberculatus* 1 and 8) were placed in the *P. chinai* clade. These samples were collected in Loja province (Fig 2.9). Because the initial classification of the bugs was conducted using morphological characters, it is possible that some samples were misidentified due to morphological variations. This dataset was not run using the ML
method implemented in GARLI because this software does not allow for data partitioning and the use of different nucleotide substitution rate models.

In addition, twelve ITS-2 sequences representing the three species were compared to other species of *Panstrongylus* found in South America. A nucleotide substitution model was tested for the dataset including the South American samples published by Marcilla et al. (2002) using jModelTest. The selected model was GTR + Γ. The nucleotide frequencies were FreqA = 0.3271, FreqC = 0.1148, FreqG = 0.1276, and FreqT = 0.4306. The gamma shape (α) = 1.0530. The resulting tree showed a clade that successfully grouped the species *P. megistus*, *P. rufotuberculatus*, and *P. herreri/lignarius*; however, the relationship among them was left unresolved (Fig 2.10). These branch splits did not present strong statistical support (posterior probability = 0.67–*P. herreri*; posterior probability = 0.61 – *P. rufotuberculatus* and *P. megistus*) using the Bayesian method. However, the ML method (Likelihood = 1) strongly supported an unresolved polytomy formed by *P. herreri/P. lignarius*, *P. rufotuberculatus* and *P. megistus*. The second clade showed *P. geniculatus* as the second group in the polytomy. The last group in the polytomy, includes *P. howardi* and *P. chinai* individuals with strong statistical support (posterior probability = 1; Likelihood = 1). No evidence of differentiation and monophyly is found between these two species using ITS-2. Therefore, this marker failed in resolving the inter-specific relationship between *P. rufotuberculatus* and *P. chinai* in Loja and Manabi provinces, Ecuador.
2.4 Discussion

Phylogenetic inference and classification of species of triatomines is of scientific interest due to the role that these insects play in the transmission of Chagas disease or American trypanosomiasis. This study has been designed to contribute to our understanding of the evolutionary history of three species of *Panstrongylus* found in Loja and Manabi provinces in Ecuador through the inference of phylogenetic trees using two genetic markers (Cyt b and ITS-2). The three species studied were *P. chinai*, *P. rufotuberculatus*, and *P. howardi*. Initially, Lent and Wygodzinsky (1979) reconstructed the phylogeny of *Panstrongylus* using morphological characters. This tree showed similar topology to the phylogenetic inference conducted using ITS-2 data by Marcilla et al. (2002). However, molecular data from both Cyt b and ITS-2 regions of mitochondrial and nuclear DNA, respectively, promise more accurate and detailed inferences revealing phylogenetic relationships. The ITS-2 nuclear marker is considered of utility to resolve relative recent divergence (< 50 million years ago) among closely related taxa, it has been used to diagnose species and infer hybridization (Mas-Coma and Bargues, 2009). On the other hand, mitochondrial DNA evolves at a faster rate than most nuclear protein-coding DNA, making it a marker suitable to study closely related species or recently diverged taxa (De Salle et al., 1987; Moriyama and Powell, 1997; Monteiro and Pierce 2001; Avise, 1994; Simon et al., 1994; Page and Holmes, 1998). Furthermore, the increased accuracy of this study compared to other studies is given by a larger dataset and the use of two molecular markers instead of one single gene to define the relationship among the studied taxa.
As shown by the results of this study, Cyt b haplotypes of *P. rufotuberculatus* were scattered within the two *P. chinai* and *P. howardi* clades reconstructed in the tree, while the phylogeny of the ITS-2 locus distinguished a clade of *P. rufotuberculatus* that was clearly divergent (range of divergence between clades = 11.7% - 18.1%) from a clade that included intermixed individuals of *P. chinai* and *P. howardi*. While one of the trees might represent the “true” phylogeny of the species, the discrepancies observed between Cyt b and ITS-2 gene trees decreased resolution to infer the species phylogeny and obtain a species tree close to this “true” phylogeny. However, incongruences between gene trees allow researchers to observe other events that form part of the evolutionary history of species. Incongruences between gene trees and the true species phylogeny can be due to a number of reasons, including hybridization, incomplete lineage sorting, retention of ancestral polymorphism, and gene duplication (Maddison, 1997). In this case, the characteristic smaller population effective size of mitochondrial DNA reduces the possibility of finding discordance between gene trees due to incomplete lineage sorting (Ballard and Rand, 2005). However, the discrepancies found between Cyt b and ITS-2 gene trees suggest that the specimens of *P.rufotuberculatus* are hybrids in which mitochondrial introgression occurred. Introgression is hypothesized when a taxon is placed within a defined clade in one gene tree, and the same taxon is found included within a different group in another gene tree (Sang and Zhong, 2000; Mas-Coma and Bargues, 2009).

The results of my study suggest that hybridization has occurred between *P. rufotuberculatus* and *P. chinai*, and between *P. rufotuberculatus* and *P. howardi*, leaving
*P. chinai* and *P. howardi* mitochondrial introgressed elements in the genome of *P. rufotuberculatus*. On one hand, the species *P. rufotuberculatus* forms a statistically supported clade when only ITS-2 sequences are compared. However, when Cyt b sequences are used, the *P. rufotuberculatus* clade is not recovered. Indeed, some members of *P. rufotuberculatus* (*P. rufotuberculatus* 1, 4, 10 and 9) are grouped within the clade of *P. chinai*, and others (*P. rufotuberculatus* 7, 2, 11, 5, and 6) are grouped within the *P. howardi* clade. Discrepancies between mitochondrial gene trees and trees of other gene markers can be the result of the higher likelihood of mitochondrial DNA to introgress, which usually involves the whole mitochondrial genome (Mas-Coma and Bargues, 2009; Ballard and Withlock, 2004). In fact, mitochondrial introgression in the apparent absence of nuclear introgression has been observed among allopatric populations (Ballard and Withlock, 2004). Moreover, levels of presumed introgression are higher in cases of sympatric or parapatric than of allopatric species (Ballard and Withlock, 2004). In this case, *P. rufotuberculatus* is found in sympatry with *P. howardi* and *P. chinai* in the Loja and Manabi provinces. Moreover, the apparent erroneous positions in Cyt b and ITS-2 trees of the haplotypes of *P. chinai* and *P. rufotuberculatus* (*P. rufotuberculatus* 1 and 8; *P. chinai* 24) might be the result of the hybridization that could induce morphological variations, which can cause their misidentification. In fact, it has been previously observed that morphological intermediates exist as product of hybridization, but their identification is only reliable when examining multiple character sets (Dowling and Secor, 1997; Mas-Coma and Bargues, 2009).
Hybridization events among triatomine (Herrera-Aguilar et al., 2009) species and the *Panstrongylus* genus (Barret, 1988) have been observed in nature as well as in the laboratory (Mazzotti and Ozorio, 1942; Perez et al., 2005; Martinez-Ibarra et al., 2009; Martinez-Hernandez et al., 2010). However, the use of more nuclear loci will be needed to rigorously test the introgression hypothesis. Moreover, additional studies on hybridization among species of *Panstrongylus* in Ecuador should be conducted through surveys of their natural habitats as well as through laboratory experimentation. It is important to establish the presence of contact zones between the species within each province by expanding the areas where samples are collected. Hybridization is of epidemiological importance because it raises the possibility that hybrids might diverge into a third species that can present enough genetic variability and acquire resistance to insecticides used in vector control, as well as improved environmental adaptability or vector competence (Mas-Coma and Bargues, 2009).

The specific relationship between *P. howardi* and *P. chinai* found in the Loja and Manabi provinces is of particular interest since the phylogeny of these two closely related species has remained unclear. The results of this study showed divergence only in the phylogeny using Cyt b haplotypes (4.3% - 8.3%) and not when ITS-2 sequences are analyzed. Incomplete lineage sorting and retention of ancestral polymorphism could account for incongruence between the gene trees. However, this possibility can be discarded because there is a higher probability that mitochondrial DNA (Cyt b) more accurately traces a short internode than ITS-2 nuclear autosomal DNA (Moore, 1995) due to the maternal nature of mitochondrial DNA and the size of its effective population, one
fourth that of nuclear loci (Moore, 1995; Ballard and Withlock, 2004). In addition, Cyt b has the highest rate of evolution among the mitochondrial DNA genes (Mas-Coma and Bargues, 2009). These characteristics of mtDNA allow ancestral polymorphisms to be more rapidly eliminated in Cyt b than in nuclear DNA (Avise, 2000), which means that mitochondrial DNA is expected to coalesce on a common ancestor more rapidly than nuclear DNA (Hudson and Coyne, 2002; Mas-Coma and Bargues, 2009). Therefore, caution is advised when using only mitochondrial DNA to define species (Moritz et al., 1992; Mas-Coma and Bargues, 2009).

Genealogical species have been defined as a group of organisms more closely related to each other and reproductively isolated (Cracraft, 1983; Page and Holmes, 1998; Mas-Coma and Bargues 2009). In addition, phylogenetic species have been defined as reciprocally monophyletic groups formed from the analysis of complete or partial loci (Mas-Coma and Bargues, 2009). The resulting tree reconstructed after using concatenated data recovered reciprocally monophyletic clades of *P. rufotuberculatus, P. howardi, and P. chinai*, where *P. rufotuberculatus* was segregated as a distinct species in monophyletic relationship with *P. howardi* and *P. chinai*, which formed sister clades.

Although these results seem to resolve the relationship between *P. howardi* and *P. chinai*, the specific resolution of conflicts between the topologies of the ITS-2 and Cyt b gene tree in this study should be carefully interpreted because other studies have found unclear criteria applied when combined data trees are resolved (Fisher-reid and Wiens, 2011). This is particularly important in this study, given that conflicting nodes of Cytb and ITS-2 are strongly supported and such discordance is best explained by introgression.
Previous studies on species of the *Panstrongylus* genus found inter-specific genetic distances ranging from 12% to 33% when ITS-2 was analyzed (Mas-Coma and Bargues, 2009). Unfortunately, no analyses of the Cyt b fragment in species of this genus have been conducted. In addition, genetic comparisons at the sub-specific level of the genus Triatoma, closely related to *Panstrongylus* found genetic distances ranging between 0% and 7%, when ITS-2 was analyzed and between 13% and 1.6% of polymorphic sites when Cyt b was studied (Mas-Coma and Bargues, 2009). These values confirm the tendency of Cyt b to change more rapidly than ITS-2 in triatomines.

Although the genetic distance values reported by Mas-Coma and Bargues (2009) were obtained as *p*-distances (raw data), they can be compared to the values obtained in this study since these, calculated using an evolutionary model, are very similar to the *p*-distances obtained from the same data. The genetic distances between *P. howardi* and *P. chinai* found in this study (0% - 2% for ITS-2; 4% - 8% for Cyt b) suggest that the level of divergence between these two groups could confer them the rank of sub-species, as previously suggested by Patterson et al. (2009). However, for the characterization of these two groups, other means of differentiation should also be taken into account.

Previously, the systematic validity of species of triatomines has taken into account ecological and melanic differences, as well as allopatric geographic distributions. *Panstrongyus howardi* and *P. chinai* represent a good example of two groups that are ecologically separated given the different habitats in which they are found. In the Loja province, *P. chinai* is mainly found in domestic and peri-domestic habitats, while *P. howardi* is found in sylvatic, domestic and peri-domestic environments in Manabi.
province. In this regard, it is also important to note that the climates of the two provinces vary in temperature, humidity, hosts, and vegetation. In addition, *P. chinai* presents a darker pigment, which constitutes a melanic differentiation, and slightly different morphology (Patterson et al., 2009). Finally, *P. chinai* and *P. howardi* have been found in separate geographic areas, suggesting the occurrence of allopatric speciation. Taking these differences, it is possible to suggest that *P. chinai* and *P. howardi* can be referred as “evolutionary units” (Dujardin et al., 2009). In fact, Dujardin et al. (1999) argue that ecological differentiation influences developmentally plastic morphological changes in populations, which can occur before genetic differentiation. Therefore, ecological factors accompanied by geographical isolation could be important mechanisms of speciation in triatomines (Mas-Coma and Bargues, 2009).

In spite of the occurrence of geographic isolation and the observation of ecological and morphological differences, reproductive isolation should be tested between these two species to validate their systematic rank of species. *Panstrongylus chinai* and *P. howardi* are currently found in two different geographic locations. However, whether zones of contact between these two groups or fertile reproduction exist is still unknown. In fact, the specific or sub-specific ranking of genetically closely related species of triatomines has relied mainly on the capacity of these different groups of producing fertile hybrids (Remigio and Blair, 1997; Usinger, 1944; Ramsey et al., 2000; Mas-Coma and Bargues, 2009) since a number of authors argue that any degree of genetic exchange can prevent speciation (Dujardin et al., 1999; Mas-Coma and Bargues, 2009).
In consequence, taking into account the differences in geographic locations, ecological habitats (P. howardi in Manabi, and P. chinai in Loja), and morphological traits (P. chinai is darker than P. howardi) between the two groups, the pattern of divergence observed in the Cyt b dataset is probably the result of time, adaptation, and geographic isolation of the two species between the Loja and Manabi provinces (allopatric speciation).

It is evident that the sequence divergence between P. chinai in Loja and P. howardi in Manabi provinces is not large, therefore it can be presumed that this event is recent. The fact that divergence is not detected by ITS-2, but is detected by Cyt b indicates the possibility of the existence of one single species that is at the border of intra- and inter-specific divergence, continuing the process of divergence. However, referring to P. chinai as a melanic subspecies of P. howardi, as stated by Patterson et al. (2009), might be inaccurate because a zone of sympatry of P. chinai and P. howardi has not been found yet; currently, because of geographical isolation, no natural hybridization or gene flow between the two groups is known to occur.

Using ITS-2 sequences, it is not possible to distinguish between P. howardi and P. chinai. Sequences of this gene lack sufficient differentiation to identify them as reciprocally monophyletic. This was also the case when samples of P. howardi, P. chinai and P. rufotuberculatus were combined with ITS-2 sequences from Marcilla et al. (2002). The gene tree reconstructed using these sequences grouped P. rufotuberculatus within its corresponding clade in the species tree. Moreover, samples of P. chinai were also grouped with another sample from the previous study of the same species. However, P.
howardi was also grouped within the P. chinai clade. The reconstructed tree did not resemble the relationships among species observed in the ITS-2 gene tree published by Marcilla et al. (2002), which included only Panstrongylus sp., particularly those relationships in which branch support was weak. Such discrepancy might be the result of the selection of a different nucleotide substitution model to conduct the analyses. The selected model for this study was based on the model that best accounted for variation in the data, as implemented in jModeltest. In addition, differences among branching patterns with low statistical support found using ML and Bayesian analyses conducted by Marcilla et al. (2002) and the present study can also account for the discrepancies. Such discrepancy, therefore, leaves the relationships among the species unclear in this tree, which supports the use of more than one genetic marker to infer phylogeny.

The use of ITS-2 to resolve phylogenetic relationships among species of Panstrongylus has been successful in some cases (Marcilla, 2002). However, the relationship between P. chinai and P. howardi in Loja and Manabí was not resolved when only ITS-2 is used, overlooking the divergence showed by Cyt b, which made both groups appear as one definite species. Lack of divergence between two groups of taxa when ITS-2 was analized has been previously interpreted as the presence of one single species, even though small values of genetic differentiation were observed in Cyt b (Martinez et al, 2006; Pfeiler et al., 2006; Monteiro et al., 1999; Bargues et al., 2006; Gumiel et al., 2003; Mas-Coma and Bargues, 2009). A concatenated data set provided better resolution of the phylogenetic relationship between these two species. Although,
the striking differences between the two gene trees, independently analyzed, indicated the necessity of using more genetic markers to infer phylogeny among these species.

This study highlights the importance of using more than one molecular marker to infer phylogeny among closely related species of triatomines. Moreover, this study also underlines the importance of considering not only genetic data but also biological information about groups of related taxa in order to evaluate their systematic classification. The use of multiple loci to conduct phylogenetic analysis would contribute to a more accurate estimation of the evolutionary history of the species, including the detection of events like hybridization, which can also cause incongruent phylogenies. In conclusion, by using two molecular markers (mitochondrial and nuclear) this study has been able to find the occurrence of introgression among species found in the Loja and Manabi province and determine the specific status of *P. howardi* and *P. chinai* as two evolutionary units diverging from one single species.
CHAPTER 3: PHYLOGEOGRAPHIC STUDY OF *PANSTRONGYLUS SP.* IN ECUADOR

3.1 Introduction

Phylogeography encompasses the study of relationships between the spatial distribution and the genealogy of species (Avise, 2000). This field uses the input from other disciplines like ethology (animal behavior), demography (changes in population structure), and population genetics (the factors that influence the genetic variation within a species) (Avise, 2000; Wakeley, 2008) in order to make inferences. The study of the relationships between the spatial distribution and the genealogy of *Panstrongylus sp.* in Loja and Manabí is of scientific importance because such information will help in the designing of effective vector control programs aiming to prevent the transmission of *T. cruzi* by triatomines found in the domicile, peri-domicile and sylvatic environments of rural communities of Ecuador (Grijalva et al., 2005).

I have analyzed the phylogeography of triatomines in the genus *Panstrongylus* (Hemiptera: Reduviidae) collected from 28 rural communities in both Loja and Manabí. Rural communities in these provinces have been subject of vector control programs, which used two-person teams (WHO, 2002) to conduct entomological searches inside and outside houses, and the subsequent spraying of pyrethroid insecticide on internal walls of the houses where bugs were found (Grijalva et al., 2011). Such searches have been conducted since 2002 to the present with constant monitoring of the houses conducted six and twelve months after the first visit. However, sometimes the yearly
monitoring of the communities has not visit the same house as the previous visit. (Grijalva et al., 2005; Grijalva et al., 2011).

Vector control and monitoring programs implemented in these areas have revealed several species of triatomines infesting and reinfesting houses, even after insecticide spraying (Grijalva et al., 2005; Grijalva et al., 2011). Moreover, in many cases, such re-infestations involve a species of triatomines different than the species present before spraying (Grijalva et al., 2011). Therefore, information about the patterns of geographical and genetic distribution of the species in these areas will help to identify potential strategies for vector control programs targeting these species.

Previous phylogeographic and biogeographic studies focused attention on two of the most important vectors of *T. cruzi* in the genera *Triatoma* and *Rhodnius*, their interspecific, intraspecific and vector-parasite relationships (Bargues et al., 2006; Abad-Franch and Monteiro, 2007; DaSilva et al., 2007; Bargues et al., 2008). The results of such studies provided valuable information about the biology of these species, allowing researchers to plan more effective vector control programs. However, because *T. cruzi* has developed the ability to use other species of triatomines as vectors, and it has been observed that the control of one species has brought reinestation of a different species (Grijalva et al., 2011), it is essential to determine the vectors historical and genetic characteristics as well.

Different triatomine species inhabit Ecuador in association with the diverse geographic and climatic nature of its regions. Loja is located at the southernmost part of Ecuador (Figure 2.1). The climate in Loja varies from tropical savanna to humid and
semi-humid. High altitudes are characterized by moorland. Dry and arid areas, especially in areas near Peru, characterize this province. Loja’s topology makes access to many places difficult. The elevations of Loja are relatively low, and the average temperature is 16°C (IGM, 1995; IGM, 2010; INAMHI, 2010). The samples were collected from communities in Loja located in forested areas at altitudes ranging from 675 to 1,423 meters above sea level (Figure 3.7).

Manabí is located in the coastal region of Ecuador (Figure 2.1), and its climate is tropical humid, reaching temperatures between 25°C to 36°C. This province has few elevations, not higher than 700 m ASL (IGM, 1995; LNS, 2009; INAMHI, 2010). The study communities in Manabí are located in forested areas at altitudes ranging from 52 to 391 meters above sea level (Figure 3.1). Loja and Manabí provinces are separated by the Andes mountain chain, at a distance of approximately 300 km from each other. Natural active dispersal of triatomines between these provinces might not be possible because commonly triatomines do not fly further than 200 m; although, it has been found that some sylvatic species can fly further than 500 m (WHO, 2002). Alternatively, passive dispersal, assisted by humans can be hypothesized (WHO, 2002). Therefore, to investigate the phylogeography of triatomines in these provinces, it is important to take into account the biological and genetic characteristic of the bugs. Such characteristics can be determined by analyzing the possible human-bug interactions in the area of study as well as analyzing genetic markers that facilitate distinguishing micro-evolutionary events in the species.
Phylogeographic studies have been conducted by using cytochrome b, ITS-1, and ITS-2 loci independently (Da Silva et al., 2007; Bargues et al., 2006; Bargues et al., 2008). Mitochondrial DNA has commonly been used for phylogeographic studies given its maternal inheritance, haploidy, and lack of recombination (Avise, 2000). Nuclear DNA is characterized by containing highly conserved loci, and has been used in macro-evolutionary studies. However, the Internal Transcribed Spacers (ITS) – 1 and – 2 are non-coding regions contained between the 18S, 5.8S and the 28S rDNA coding regions. These spacers are characterized by a faster rate of evolution than the coding regions, which makes them suitable for micro-evolutionary studies (Marcilla et al, 2002). Several phylogeographic studies of triatomines have been conducted using ITS-2 (Bargues et al., 2006; Bargues et al., 2008; Dorn et al., 2009; Quisberth et al. 2011), which demonstrates the suitability of this marker for such analyses.

Genetic diversity is one of the first fundamental concepts studied in population genetics (Templeton, 2006). In fact, it has been found that genetic diversity influences the “fitness” of a population. High diversity within a population is related to its survival, while low diversity is related to its extinction under environmental changes or stress (Markert et al. 2010). The sources of diversity within a population are mutations, genetic drift, migration, and selection. Genetic diversity can be measured from several kinds of data, using several calculations (Hartl, 2000; Templeton, 2006; Wakeley, 2008). Using sequenced DNA data, genetic diversity within a population can be measured in function of the polymorphisms present in a sample ($\pi$), and the haplotype variability among a group of collected samples ($h$).
The demographic history of a species is of fundamental importance in phylogeographic studies because of the impact that demographic changes have on the genetic structure of the species (Avise, 2000). The coalescent model in population genetics assumes the neutral evolution of the gene trees and is used for interpreting DNA sequence data in the context of population genetics (Ramos-Onsins and Rozas, 2002). Several authors have found that demographic changes in a population leave their fingerprint in the DNA sequence data, which can be detected by comparing the pairwise differences distributions among sequences (Tajima, 1989; Slatkin and Hudson, 1991; Rogers and Harpending, 1992). Other authors have also found that using information about the mutation frequency found in the sequences, and haplotype distribution can also serve to detect demographic changes within a population (Ramos-Onsins and Rozas, 2002; Fu, 1996). Four tests (Tajima’s $D$, Fu’s $F$, $R_2$ and Mismatch distributions tests) were conducted in this study in order to detect demographic events within each species and evolutionary unit of *Panstrongylus* found in Loja and Manabí.

The genetic structure of a species and the degree of differences among individuals within and among populations are influenced by their ability to disperse and mate with members of other populations (Slatkin, 1993). According to Wright (1943), populations continuously distributed can become isolated -by- distance, when the species has short-range means of dispersal. According to the isolation- by- distance model, it is expected that individuals from localities close to one another have more genetic similarity than those located more distantly apart (Rousset, 2004). Isolation by distance can be tested using the Mantel’s test, which compares a matrix of geographic distances to a matrix of
genetic distances (Sokal and Wartenberg, 1983; Sokal and Rohlf, 1995). Genetic distances calculated using Wright’s $F_{ST}$ values can be used to conduct the Mantel’s test. Wright’s $F_{ST}$ is used as a measure of the random differences between subpopulations, where $F_{ST}$ will increase as the similarity of genes within subpopulations increase and more differences are found between populations (Rousset, 2004).

3.2 Materials and Methods

3.2.1 Genetic diversity indices

To analyze the level of genetic diversity among samples within each species and evolutionary unit of Panstrongylus found in Loja and Manabí provinces, the nucleotide diversity index ($\pi$), haplotype diversity index ($h$), number of segregating sites in a sample ($S$), and number of haplotypes ($H$) were computed within each species and evolutionary unit. Calculations were conducted using Arlequin ver. 3.5.1.2 (Excoffier and Lischner, 2010) and DNAsp ver. 5.10.01 (Librado and Rozas, 2009). Genetic diversity indices were independently calculated using sequences from Cyt b and ITS-2 for each species and evolutionary unit of Panstrongylus. The three groups of Panstrongylus were analyzed independently.

Intra-population diversity was analyzed using 11 sequences from samples of P. howardi collected in El Bejuco and 8 sequences from samples of the same evolutionary unit collected in La Cienega communities in Manabí province. Analyzed sequences included Cyt b and ITS-2 markers. Genetic diversity indices were calculated within these
two communities only, due to insufficient sampling in the rest of the communities. For these calculations, total number of mutations was calculated.

3.2.2 Intra-specific genetic structure

To study the degree of genetic differentiation among samples collected in different geographical locations (communities) in Manabí and Loja provinces, haplotypes among communities were compared to each other. For this purpose, the fixation index \( F_{ST} \) was calculated using Arlequin. Calculations were set up to compute an \( F_{ST} \) matrix using pair wise differences. Significance was tested by conducting 1,000 permutations. Cyt b and ITS-2 sequences were independently tested for this purpose. Additionally, microsatellites in ITS-2 sequences were examined to test for correlations between tandem repeats and sample locations (communities) within each species and evolutionary unit.

3.2.3 Haplotype networks

Haplotype networks were constructed to analyze the relationship among haplotypes found within each species and evolutionary unit of Panstrongylus using TCS ver. 1.21 (Clement et al., 2000). Prior to the analysis, Cyt b and ITS-2 sequences were truncated to an equal length within each alignment. Haplotype networks were calculated using a statistical parsimony algorithm, which only links haplotypes that can be connected with a 95% confidence (Templeton et al., 1992).
3.2.4 Demographic History

Neutrality tests were conducted to determine whether haplotypes from each of the three species and evolutionary unit of Panstrongylus were in mutation-genetic drift equilibrium, according to the neutral theory expectations (Kimura, 1983). Demographic history was analyzed using four tests that use different information from the DNA sequences. Tajima’s $D$ and Ramos-Onsins Rozas ($R_2$) use mutation frequencies in the sequences to determine if a population has undergone a recent population expansion event. Tajima’s $D$ is based on the differences between nucleotide diversity and segregating sites; while $R_2$ is based on the comparison of the number of singletons on a genealogy branch after a sudden population expansion and a stable population (Ramos-Onsins and Rozas, 2002; Hurwood et al., 2008). Positive values of Tajima’s $D$ are associated with balancing selection where frequencies of polymorphisms are too equal, and negative values are associated to selective sweeps or population expansion after a recent bottleneck where polymorphisms are too unequal. Lower values of $R_2$ are expected under the population expansion model (Tajima, 1989; Hartl, 2000; Hurwood et al., 2008, Ramos-Onsins and Rozas, 2002).

Fu’s $F_S$ test uses information from the haplotype distribution in a sample. The test estimates the probability of observing a random sample with equal or less singletons than the observed given a level of diversity. The test is based on the infinite site’s model of mutation, and assumes that all of the alleles are selectively neutral. Negative values of $F_S$ will be caused by excess in singletons given a population expansion event, or from genetic hitchhiking. Positive values of $F_S$ indicate a deficiency of singletons, which is
expected from a recent population bottleneck or over dominant selection (Ramos-Onsins and Rozas, 2002; Joyce et al. 2003; Holsinger, 2010).

Finally, given that population expansion events leave a characteristic distribution of pairwise sequence differences, this information can also be used to test demographic changes. However, this test statistic should be conservative in terms of the \( p \)-value used to accept or reject the null hypothesis of demographic expansion (Kuchta et al., 2009).

Fu’s \( F_S, R_2 \), and mismatch distribution tests were independently conducted on Cyt b and ITS-2 sequences of 3 species groups of Panstrongylus. Tajima’s \( D \) test was conducted only on Cyt b sequences since ITS-2 sequences are not protein-encoding sequences. This test was conducted using Cyt b fragments aligned, truncated to 646 bp, and coded according to the \( P. megistus \) Cyt b nucleotide and protein sequences (gi|5051600|gb|AF045722.1|_Panstrongylus_megistus_cytochrome_b_(cytb)_gene,_mitochondrial_gene_encoding_mitochondrial_protein,_partial_cds) published by Lyman et al. (1999). Computations and permutations for each test were conducted using Arlequin and DNAsp software, where segregating sites were considered. The number of permutations was set to 10,000 for all the tests, and \( p \)-values were calculated for a 95% confidence interval.

3.2.5 Geographic dispersal

The Mantel’s test incorporated in the software IBDWS (Bohonak, 2002; Ngan, 2006) was used to determine whether there is a correlation between genetic and geographic distances among populations. According to the isolation-by-distance model,
geographical distance is proportional to genetic distance among sampled individuals (Wright, 1943). The calculations were set up using 10,000 randomizations using \( \Phi_{ST} \) genetic distances between populations calculated using the Kimura-2-parameter model (K2P) and linear geographic distances among these populations. Gaps were treated as transversions. The software also used the Rousset’s distance measure (Rousset, 1997) to analyze population subdivision according to the IBD (Isolation-by-distance) model. Finally, the slope, y-intercept, and the coefficient of determination \( (r^2) \) of the IBD plots were calculated using the RMA (Reduced Major Axis) regression (Bohonak, 2002), also incorporated in the software.

The first analyses were conducted using all the sequences of samples collected within each species and evolutionary unit of \textit{Panstrongylus} to observe the general trend in genetic differentiation. Subsequent analyses were conducted using specific linearly dispersed populations within a geographical area (i.e. province). In addition, samples from clustered locations within such areas were analyzed. These different analyses were conducted to observe if clustered samples followed the same dispersion and genetic patterns as dispersed samples.

3.3 Results

3.3.1 \textit{Panstrongylus howardi}

3.3.1.1 Genetic Diversity

\textit{Panstrongylus howardi} samples were collected from rural communities located in the central region of Manabí within an area of approximately 2,000 km\(^2\) (Fig 3.1). Within
this area, the species possesses the lowest levels of genetic diversity, compared to *P. chinai* and *P. rufotuberculatus*, indicated by lower nucleotide and haplotype diversities for both Cyt b and ITS-2 sequences (Table 3.1). Furthermore, low intra-population diversity was observed within El Bejuco and La Cienega communities in Manabí province, as shown in Table 3.1.

### 3.3.1.2 Intra-specific structure

Fixation indices were calculated to measure the genetic distances among haplotypes of Cyt b and ITS-2 from rural communities in Manabí province, where *P. howardi* was collected. No significant differences were found among most of these communities when Cyt b sequences were used (Table 3.2). Two significant differences were found between La Cienega and Zapallo, and between La Cienega and San Gabriel communities (*p* < 0.05). When ITS-2 sequences were compared, no significant differences were found among most communities, supporting previous calculations, which indicate the prevalent homogeneous nature of the *P. howardi* haplotypes in Manabí. However, significant differences were found between the Zapallo and San Gabriel, and between El Bejuco and San Gabriel (*p* < 0.05), in spite of being geographically closely located (Table 3.3; Fig. 3.1). Negative $F_{ST}$ values can be interpreted as no differences between the two populations compared, due to imprecision of the algorithm used (Jaramillo et al., 2001).

The complex microsatellite AATTT(AT)$_5$AGTTTATT(AT)$_6$ was found within most of the ITS-2 sequences of *P. howardi*. Two variants of this microsatellite, found
only in two different geographic locations differ in the number of AT repetitions (5 and 7 repetitions) on the last segment of the microsatellite \((AT)_6\). Other sequences within the species and evolutionary unit were characterized by the absence of the complex microsatellite and the presence of partial sections of it. Microsatellites were not characteristic of a particular population or habitat. Therefore, microsatellites cannot be used to identify individuals according to population or habitat.

3.3.1.3 Haplotype networks

Haplotype networks constructed using Cyt b confirmed the low genetic diversity among \(P. howardi\) in populations in Manabí by showing that most of the samples collected in El Bejuco, La Cienega, Zapallo, Santa Rosa de las Palmas, La Encantada and Maconta Abajo shared a single haplotype (Fig. 3.2B). The genetically most distant haplotype corresponds to one sample collected in Pacoche; the second genetically most distant haplotype comes from a sample that was collected in El Bejuco. Other haplotypes also distributed within the same populations varied from 1 to 10 nucleotide positions relative to the most common haplotype (Fig. 3.2).

Haplotype number increased to 25 within \(P. howardi\) when ITS-2 sequences were used. Two haplotype networks and an isolated haplotype resulted from the network reconstruction with a 95% confidence level. Most haplotypes (22) are distributed in all the sampled populations, and differ from 1 to 9 nucleotide differences (Fig 3.3A). Networks were resolved as separated when more than 9 nucleotide differences were found. The second network included 3 different haplotypes, which were also distributed
through the central area of the province. Finally, one isolated haplotype was found in La Cienega (Fig 3.3A). The geographical distribution of these haplotypes within Manabí province was characterized by the presence of 3 to 6 haplotypes found within one single community. However, cases of shared haplotypes among different communities can also be found (Fig 3.3B).

3.3.1.4 Demographic history within *P. howardi*

Demographic changes within *P. howardi* in Manabí province were tested using Fu’s *F*<sub>S</sub>, Ramos-Rozas *R*<sub>2</sub>, and mismatch distribution using Cyt b and ITS-2; and Tajima’s *D*, using Cyt b. The *F*<sub>S</sub> value for Cyt b was significantly negative (*p*-value < 0.05), suggesting a recent demographic expansion. In addition, *R*<sub>2</sub> was significant, also suggesting recent population expansion of *P. howardi* in Manabí province (Table 3.4). Tajima’s *D* was also significantly negative, supporting population expansion, which might have followed a recent population “bottleneck” with insufficient time to restore the equilibrium between mutation and random genetic drift. Besides this, the negative value of Tajima’s *D* also supports demographic expansion rather than balancing selection. Finally, mismatch distributions showed a pattern consistent with population expansion (Fig 3.4A). In addition, ITS-2 sequences were also used to conduct the tests. In support of the previous data, significant values of Tajima’s *D*, *F*<sub>S</sub>, *R*<sub>2</sub> (Table 3.4), and patterns of mismatch distribution (Fig 3.4B) evidenced recent population expansion among *P. howardi* in Manabí.
3.3.1.5 Geographical Dispersal of *P. howardi* within Manabí

Isolation by distance using Cyt b sequences to compare all the populations in Manabí was not significant (*p*-value = 0.06) for *P. howardi* (Fig 3.5A). However, when the test was conducted using the samples collected in La Cienega, Maconta Abajo, Zapallo, and Pacoche, isolation by distance was significant (*r* = 0.98; *p* < 0.001; Fig 3.5B). These communities are distributed within a linear range of approximately 73 km within Manabí province (Fig 3.1). Additionally, a third test of IBD was conducted using sequences from the communities El Bejuco, San Gabriel, Zapallo, and La Encantada, which was not significant (*r* = 0.4418; *p* > 0.1; Fig 3.5C). The results of these tests suggest that there is genetic similarity among communities located within a range of 10 km, and that geographical distance correlates with genetic distance among communities located out of such range (Fig 3.3B). When the same tests were conducted using ITS-2 sequences, IBD was not found among all the populations (Fig 3.6A), or among La Cienega, Maconta Abajo, Zapallo and Pacoche populations (Fig 3.6B).

3.3.2 *Panstrongylus chinai*

3.3.2.1 Genetic Diversity

*P. chinai* samples were collected from rural communities distributed through the central region of Loja province within an area of approximately 3,800 km² (Fig 3.7). This evolutionary unit presents lower Cyt b nucleotide and haplotype diversity indices than *P. rufotuberculatus*, but similar values to *P. howardi* (Table 3.1). Using the ITS-2 marker, haplotype diversity, haplotype number and nucleotide diversity are increased (Table 3.1).
Such increased variability found in ITS-2 sequences is expected, given the diploid nature of this marker.

### 3.3.2.2 Intraspecific Structure

As previously described, *P. chinai* sequences were used to calculate $F_{ST}$ to measure the genetic differentiation among samples collected in 12 populations in Loja province. Fixation indices were calculated separately for Cyt b and ITS-2 sequences. Cytochrome b haplotypes from populations in Loja province exhibited no significant differentiation (Table 3.5). In addition, when ITS-2 sequences from the populations were compared no significant difference was found (Table 3.6). These comparisons indicate that the populations of *P. chinai* in Loja province are genetically homogeneous and unstructured.

Analysis of microsatellites within *P. chinai* showed the presence of the complex microsatellite AAATTTT$(AT)_5$AGTTTATT$(AT)_6$ in all the samples of Santa Ester, Chaquizca, La Extensa, and Jaguay. This complex microsatellite was also found in samples from Chirimoyos, Guara, and Naranjo Dulce. However, in these populations, the genotype AATT was also found replacing the complex microsatellite. Samples collected in Ashimingo presented three different genotypes where the complex microsatellite, the AATT genotype, and the total absence of the microsatellite were observed. Moreover, San Francisco was characterized by the presence of one of the variants of this microsatellite (AATT$(AT)_3$TTT$(AT)_5$). Finally, Lucarqui and Coamine are also characterized by the presence of the genotype AATT. These observations indicate that
there might be gene flow occurring among populations. For this reason, microsatellites cannot be used to identify individuals from a determined community or habitat.

3.3.2.3 Haplotype networks

With a 95% confidence interval, TCS constructed one haplotype network using Cyt b sequences. This network is formed by 10 haplotypes, which vary by 1-10 base pairs. One haplotype (H10) is found distributed among 5 communities (Chaquizca, Guara, Chirimoyos, La Extensa, and Santa Ester), which are located in the eastern region of the sampled area (Fig 3.8B). In addition, a second haplotype (H2) was found distributed among Ashimingo, Bramadero and Naranjo Dulce (Fig 3.8B), which are geographically located in close proximity. The other haplotypes in the Cyt b network were found as unique in one population, or sharing the same population with other haplotypes. Overall, *P. chinai* Cyt b haplotypes were genetically similar among each other (Fig 3.8A).

When TCS constructed networks with a 95% confidence level using ITS-2 sequences, 17 haplotypes were found. These formed two networks and three significantly different unique haplotypes (Fig 3.9A). The first network grouped haplotypes from samples collected in La Extensa, Guara, Chaquizca, and Jaguay. The second network grouped haplotypes from the samples collected in Chirimoyos, Naranjo Dulce, and Coamine (Fig 3.9A). The geographical distribution of ITS-2 haplotypes in the Loja province was characterized by the presence of single haplotypes distributed in different
populations. Although, La Extensa and Ashimingo were characterized by the presence of 2 and 3 haplotypes (Fig 3.9B).

### 3.3.2.4 Demographic history of P. chinai

Neutrality tests conducted to observe demographic changes using Cyt b showed low values for $F_S$, $R_2$, and negative values of Tajima’s $D$, although none was significant ($p > 0.1$) (Table 3.4). Similarly, mismatch distributions (Fig 3.10A) were not significant (SSD $p$-value > 0.2), which suggests that *P. chinai* in Loja province is experiencing demographic expansion. As with Cyt, ITS-2 $F_S$ and $R_2$ values were not significant (Table 3.4). In addition, population expansion of *P. chinai* was supported by a non-significant mismatch distribution (SSD $p$-value > 0.4) where differences between the expected and the observed distributions for the population expansion model were not found (Fig 3.10B).

### 3.3.2.5 Geographical Dispersal of P. chinai within Loja province

Isolation by distance was tested using Cyt b sequences from *P. chinai* samples collected in every community in the Loja province. The first test was conducted using all the samples collected in Loja province. These results showed that a weak ($r^2 = 0.24$), but significant ($p < 0.01$) isolation by distance exists (Fig 3.11A). Samples from La Extensa, Naranjo Dulce, Ashimingo, Bramaderos and Santa Ester, populations located within a linear range of approximately 50 km, were not isolated by distance (Fig 3.11B). Furthermore, when samples from Santa Ester, Chirimoyos, Chaquizca, and Guara,
located within a linear range of 15 km, were tested, no significant isolation-by-distance was found (p > 0.1; Fig 3.11C). A third tests was conducted using all the samples collected in Loja province, and it was found that, overall, isolation-by-distance was not evidenced among \textit{P. chinai}. Such apparent discrepancy might be the result of a recent structuration of populations located at further geographical distance, and the lack of structure of populations located in close geographical proximity. Using ITS-2 sequences, no significant correlation was found neither when all samples collected within Loja province were analyzed (Fig 3.12A) nor when La Extensa, Naranjo Dulce, Ashimingo, Bramaderos, and Jaguay were tested (Fig 3.12B).

\textbf{3.3.3 \textit{Panstrongylus rufotuberculatus}}

\textbf{3.3.3.1 Genetic Diversity}

Samples of \textit{P. rufotuberculatus} were collected from rural communities in Loja and Manabí. In Loja, samples were collected from 3 populations located within a range of approximately 900 km$^2$ (Fig 3.7). In Manabí, samples were collected from 9 communities distributed within an area of approximately 2,250 km$^2$ (Fig 3.1). This species presented higher haplotype diversity indices than \textit{P. chinai} and \textit{P. howardi} when samples from Loja and Manabí were compared. Furthermore, when samples were split into two groups, each corresponding to the province where they were collected, haplotype and nucleotide diversity indices in Manabí were similar to those values found when both provinces were compared (Table 3.1). Samples collected in Loja province were not tested, given the small sample size (n = 4).
3.3.3.2 Intra-specific structure

Fixation indices ($F_{ST}$) were calculated to establish genetic differences among *P. rufotuberculatus* populations collected in Loja and Manabí, using Cyt b and ITS-2 sequences. When Cyt b was used for comparisons, no significant differences were found among populations from Loja and Manabí, (Table 3.7). When ITS-2 sequences from different communities were compared, there were no significant differences either (Table 3.8). These results suggest that populations within the two geographical locations (Loja and Manabí) are genetically similar, and no genetic structure can be identified.

When ITS-2 sequences of *P. rufotuberculatus* were analyzed, two variants of one complex microsatellite were observed. The variant AAATTTT(AT)$_4$T(AT)$_6$ was found in El Bejuco, Punta Larga, Humedad and Estero Seco. The second variant AAATTTT(AT)$_3$TTT(AT)$_5$ was found in Algarobillo, Pitayo, Estero Seco, and San Francisco. In addition, the variant (AAATT(AT)$_4$TTT(AT)$_3$ was found in San Ramon. However, the genotypes AATT in El Sauce and AATTTT(AT)$_4$AGTTTATT(AT)$_6$ in Algarobillo were also found. These two genotypes were also found among some *P. chinai* and *P. howardi* samples. In spite of the apparent variability among the genotypes, these microsatellites cannot be directly associated to a specific population because the sample size in every community is too small (n = 1-2). The results of this analysis are suitable to show that microsatellites in ITS-2 sequences can be used to establish interspecific differences between *P. rufotuberculatus*, and the other two evolutionary units *P. chinai* and *P. howardi*. However, differentiation between *P. chinai* and *P. howardi* cannot be established using these microsatellites.
3.3.3.3 Haplotype networks

All *P. rufotuberculatus* samples were pooled to construct a haplotype network using Cyt b sequences. With a 95% confidence level, two haplotype networks were produced. Both networks linked together samples collected in Loja (Fig 3.13A) and Manabí (Fig 3.13B). Networks were unlinked when more than 9 nucleotide differences were found among sequences (Fig 3.13C).

Moreover, when ITS-2 sequences were used to construct haplotype networks of *P. rufotuberculatus*, one network was produced, in addition to three unlinked haplotypes. The network shows two easily distinguished groups: the first linking haplotypes H4 and H5 found in El Bejuco, Humedad, Punta Larga and Estero Seco, Manabí, located within a range of approximately 18 km² (Fig 3.14B). The second group linked H6, H7, and H8 found in Algarobillo, San Francisco, Pitayo, in Loja (Fig 3.14 A) and Estero Seco in Manabí (Fig 3.14C). Three samples collected from Algarobillo and El Sauce (Loja) and San Ramon (Manabí) are shown as unique haplotypes significantly distant from each other and from the network (Fig 3.14C).

3.3.3.4 Demographic history within *P. rufotuberculatus*

When neutrality tests were conducted using Cyt b sequences of *P. rufotuberculatus* in Loja and Manabí, $F_S$, and $R_2$ indicated demographic stability ($p > 0.1$), and Tajima’s $D$ supported demographic stability under equilibrium conditions ($p > 0.1$) of the species (Table 3.4). In support of these data, mismatch distributions conducted using all the samples collected in Loja and Manabí (Fig 3.15A) and those collected in
Manabí only (Fig 3.15B) departed from the population expansion model. Furthermore, ITS-2 samples supported demographic stability with no significant $F_S$ and $R^2$ values (Table 3.4). Mismatch distributions showing departure from demographic expansion model were obtained when samples collected in Loja and Manabí (Fig 3.16A) and those collected in Manabí only (Fig 3.16B) were tested. Samples from Manabí were tested to determine whether or not samples collected within one province displayed the same pattern of mismatch distribution. However, samples collected in Loja province could not be tested due to the small sample size ($n = 4$).

3.3.3.5 Geographical Dispersal of *P. rufotuberculatus* in Loja and Manabí.

Isolation by distance was tested using samples collected in Loja and Manabí to determine whether or not the geographical distance between samples collected in the two provinces was correlated with genetic distance. Isolation-by-distance was not found within *P. rufotuberculatus* from Loja and Manabí using Cyt b ($p > 0.1$) (Fig 3.17A). In addition, a second test was conducted to observe whether the geographical distances among samples collected in linearly located communities within the Manabí were correlated with their genetic distances. Samples from Estero Seco, San Ramon and San Francisco in the Manabí, which were located within a range of 107 km, showed no isolation by distance ($p > 0.1$; Fig 3.17B). Furthermore, when all ITS-2 sequences from Loja and Manabí were tested, significant isolation-by-distance ($p < 0.001$) was found (Fig 3.18A). However, when Manabí samples where independently tested, no isolation by distance was found ($p > 0.1$; Fig 3.18B). Samples from Loja province were not tested
independently because of the low sample size (n=4 samples). The results obtained from the testing of Cyt b and ITS-2 sequences suggest that genetic isolation might be a very recent and ongoing event between *P. rufotuberculatus* in Loja and Manabi provinces, or that populations may have been recently introduced from one place to the other.

**3.4 Discussion**

3.4.1 Phylogeography of *Panstrongylus howardi*

3.4.1.1 Genetic diversity and demography of *P. howardi* in the Manabi province

*Panstrongylus howardi* is one of the triatomine evolutionary units found infesting peridomicle, domicile and sylvatic environments in Manabi (Grijalva et al, 2005; Suarez-Davalos et al, 2010; Grijalva et al, 2011). Insecticide spraying of infested domicile areas was practiced since 2004 through 2010 in this province. Such spraying involves treating only infested houses following the entomological search (Grijalva et al, 2011). *Panstrongylus howardi* has been found reinfesting previously treated domicile areas where other species were found before insecticide treatment. The results of this study showed low indices of genetic diversity of mitochondrial and nuclear DNA markers. Among 33 samples, 12 Cyt b haplotypes and 25 ITS-2 haplotypes were found. The increased number of haplotypes in ITS-2 samples is expected given the diploid nature of this marker, compared to the haploid Cyt b. The nucleotide diversity of these samples reflected low genetic variability among haplotypes of Cyt b and ITS-2 because of the low probability of finding two different random homologous nucleotides ($\pi = 0.00292$ and $\pi = 0.00572$ respectively for both data sets). Such genetic homogeneity
among *P. howardi* might indicate that these bugs have been recently established in these areas, after spraying in the province.

Neutrality tests showed that *P. howardi* in Manabí is not in mutation-drift equilibrium. In fact, Ramons-Onzins $R_2$, and Fu’s $F_S$ tests, based on the mutation frequencies and haplotype distribution, respectively (Ramos-Onsins and Rozas, 2002) were significantly low. These results were found for both Cyt b and ITS-2 independently, providing stronger support for the demographic inferences. Mismatch distribution analyses were also conducted to illustrate the distribution of sequence differences (Ramos-Onsins and Rozas, 2002), instead of looking only at the mutational summary statistics. The results showed an excess of low frequency variants among *P. howardi* Cyt b sequences. Finally, Tajima’s $D$ value was significantly negative, which indicates the presence of low frequency nucleotide polymorphisms.

The nucleotide polymorphism patterns discussed above agree in indicating population expansion. Populations that have undergone a catastrophic event where genetic variability has been lost tend to rapidly expand to recover in number; however, they also tend to retain low genetic variability (Hartl, 2000).

The hypothesis of a recent establishment of the species in the sampled localities is supported by the neutrality tests, which indicated population expansion in both Cyt b and ITS-2 markers (Table 3.4). Therefore, it is possible that insecticide spraying in Manabi province controlled species that were competing with *P. howardi* for domiciliary habitats, allowing them to move in from sylvatic habitats and establish new infestations. In addition, applying insecticide only to infested houses might decrease the number of
individuals found in domestic and peri-domestic areas within a community. This would allow small sylvatic populations of *P. howardi* to infest or be attracted to houses and/or peri-domestic areas that have not received insecticide treatment. As reported by Grijalva et al. (2011), houses where no bugs were found at the beginning of the study were found to be infested after 6 months of the first visit. Furthermore, houses that were not infested after 6 months were found to be infested after 12 months, although these reinfestations were observed in a low percentage (Grijalva et al, 2011). For example, *P. howardi* was first found in one house of the El Bejuco community in 2007. Two years following insecticide treatment, the same house was reinfested with a bug (*P. howardi*) that presented different Cyt b and ITS-2 haplotypes, but was closely related to the first bug found in 2007. Similarly, in the La Encantada community, infestation and reinfestation in one single house was observed for two years by bugs carrying the same Cyt b haplotype. This indicates that bugs found reinfesting houses share their maternal lineage and are closely related to bugs that might have survived insecticide spraying or belong to the same genetic source. In addition, the sylvatic sample found in El Bejuco possessed the same Cyt b haplotype as most common found in most of the communities.

To conclude, insecticide spraying of infested houses or “target spraying” might not be sufficient to control triatomine infestation within certain communities because after the treatment, untreated houses are left susceptible. Thus, it is possible that some bugs escape the control intervention and move to untreated houses in the vicinity. Therefore, community interventions that can eliminate bugs from infested houses and, at the same time, limit the possibility of new infestations from sylvatic populations might be
more suitable. Taking into account population expansion and low genetic diversity, it is possible to hypothesize that re-infestation events are due to the presence of new bugs from sylvatic environments in neighboring areas, migratory bugs from one house to another, or bugs brought by humans from neighboring communities, but not due to resistance to the insecticide. Risk of acquisition of such resistance would be reflected by increased genetic variability (Dorn et al., 2009) within *P. howardi*.

### 3.4.1.2 Genetic structure and geographic dispersal of *P. howardi* in the Manabi province

$F_{ST}$ values were calculated to measure the degree of genetic differentiation among populations of *P. howardi* within Manabi. The lack of significant genetic differentiation among these populations suggests the existence of shared characteristics like habitat preference, mating behavior, tendencies toward domestication, and time and conditions of dispersal that, according to Dorn et al. (2009) might involve vector competence. More importantly, populations undergoing expansion show an excess of singletons due to recent mutations (Ramos-Onsins and Rozas, 2002), which indicates high genetic similarity among sequences. This similarity suggests that bugs from different communities might be exchanging genetic information. In terms of genetic structure, Cyt b and ITS-2 markers behaved similarly, mostly resulting in non-significant $F_{ST}$ values. However, when Cyt b sequences were tested, two significant $F_{ST}$ values among communities were found; the first between La Cienega and Zapallo (18.4 km) communities ($F_{ST} = 0.25466$ $p$-value $= 0.00586$), and the second, between La Cienega and
San Gabriel (18.4 km) communities ($F_{ST} = 1$ $p$-value = 0.01855). The genetic differences between Cyt b and ITS-2 genetic differences between La Cienega and San Gabriel might indicate higher male dispersal between these populations. It is possible that males of this evolutionary unit disperse more than females in their search for mates. Furthermore, the unstructured distribution of different ITS-2 microsatellite genotypes among the communities within Manabí supports the hypothesis of gene flow among communities. The constructed ITS-2 haplotype network showed a wide distribution of haplotypes among the studied populations, supporting the previous hypothesis that there is a high dispersal of the bugs in this area (Fig 3.3). Genetic similarity could be due to recent colonization or gene flow; however, gene flow does not necessarily follow colonization.

It has been reported that triatomines can actively move by flying or walking, as well as being passively transported by humans (WHO, 2002; Cerere et al. 2004). However, dispersal of triatomines from one community to another may be attributed to human intervention due to the long distances that exist between communities, taking into account that few species of triatomines can fly longer than 500 m (WHO, 2002). In addition, dispersal of the bugs may also be attributed to generational movement. Dispersal of triatomines among houses or from sylvatic habitats to domiciliary or peri-domestic habitats within a community might be attributed to both insect mobility and human transport. It is possible that people, frequently visit other houses with different purposes, and transport the bugs on clothing, products, animals, or other materials.

The shape and distribution of haplotypes within the network constructed using Cyt b sequences of $P. howardi$ are consistent with a population expansion model. In fact,
The parsimony method used to observe the relationships among haplotypes shows the presence of an excess of low frequency polymorphisms among the different haplotypes. The construction of haplotype networks using Cyt b and ITS-2 sequences were also helpful to illustrate the degree of genetic similarity among haplotypes collected from different geographical locations within Manabí province. It is clear that different haplotypes are distributed across the localities in the province without structure at the community level.

Isolation by distance tests, conducted on Cyt b sequences, support the hypothesis that the proportion of genetic distances correlates with geographic distances among communities of Manabí province. These results were observed when populations located within a linear range of more than 10 km were tested (Fig 3.7). For example, when comparing Cyt b sequences among the La Cienega, Maconta Abajo, Zapallo and Pacoche communities, a significant correlation was found. Nevertheless, when communities geographically located within a linear range of 10 km were compared, no correlation was found. When ITS-2 sequences were used to conduct the same tests, no correlation was found, adding more evidence to support the hypothesis that male dispersal is higher than female dispersal.

The results of the geographic and genetic distance correlations using Cyt b indicate that samples located within a range of 10 km are similar, while samples located further than 10 km apart are somewhat independent. When comparing IBD and $F_{ST}$ results, the data suggest that *P. howardi* best fits a population expansion model.
However, geographically distant populations, might be relatively demographically independent.

Such pattern suggests that gene flow might be stronger when samples are closer. Communities located within a range of 10 km experience more gene flow than communities located more distantly. Therefore, communities located within 10 km of each other are at higher risk of being re-infested with bugs of similar genetic make up (Dorn et al., 2009). Gene flow can be attributed to either human intervention, when dwellers serve as transportation to the bugs, or the proximity of sylvatic populations located in the vicinities of houses and can serve as “genetic bridges.” These populations were not sampled in this study.

3.4.2 Phylogeography of Panstrongylus chinai

3.4.2.1 Genetic diversity and demography of P. chinai in Loja province

Loja, as well as Manabi, has been part of the Chagas vector control program being carried out in Ecuador since 2005. The method of spraying only “target” domestic and peri-domestic habitats where bugs were found has been practiced in this region (Ocaña-Mayorga et al., 2010). Entomological searches that result in the finding and capture of triatomines are followed by insecticide spraying. The searches are repeated 6 and 12 months after the first visit (Grijalva, unpublished information).

*Panstrongylus chinai* is one of the triatomaine evolutionary units found in domestic and peri-domestic but not in sylvatic habitats in the Loja province (Grijalva and Villacis, 2009). This evolutionary unit has been found in synantrophy with species of the
genus *Rhodnius*, which have apparently adapted to sylvatic habitats (Grijalva et al., 2005; Grijalva and Villacis, 2009). Species other than *P. chinai* such as *P. rufotuberculatus* and *Triatoma carrioni*, can only be found in domestic and peri-domestic but not in sylvatic habitats (Grijalva and Villacis, 2009). Consistent insecticide treatment might be decreasing genetic diversity among *P. chinai* in Loja province. This study shows low genetic diversity indices among *P. chinai* for Cyt b and ITS-2 sequences. Of 26 samples, 9 Cyt b and 17 ITS-2 haplotypes were found (Table 3.1). Nucleotide diversity in *P. chinai* was similar to that observed among *P. howardi* populations ($\pi = 0.00339$). Bugs re-infesting houses one year after entomological search and insecticide treatment were found to possess the same Cyt b haplotype as the original samples. Houses found to be reinfested in the communities Bramaderos and Jaguay correspond to such pattern. The maintenance of a single Cyt b haplotype in reinfested houses might reflect incomplete elimination of the bugs after insecticide spraying as well as to the re-infestation of the house by bugs coming from exactly the same biological source as the previous one. In addition, genetic diversity within *P. howardi* was much higher ($\pi = 0.01493$) when ITS-2 sequences were analyzed. The increased heterogeneity among these samples is expected, given the diploid nature of this marker.

The null hypothesis of population stability was tested on Cyt b using the Ramos-Onzins $R_2$ and Fu’s $F_S$, as well as the null hypotheses of population expansion and neutral evolution were tested using mismatch distributions and Tajima’s $D$ respectively. In spite of obtaining values of $R_2$ and $F_S$ that best fit a population expansion model (Table 3.4), no significant results were found. In addition, when Cyt b was tested using Tajima’s $D$,
no evidence of expansion was found because a negative value was obtained, but no significance was obtained. On the contrary, ITS-2 sequences tested using Tajima’s $D$ resulted in a significantly negative value ($p < 0.05$), supporting population expansion. In addition, mismatch distributions of Cyt b and ITS-2 supported the population expansion model among *P. chinai*. Tajima’s $D$ is a test based on the differences between nucleotide polymorphisms and nucleotide diversity (Hartl, 2000), and mismatch distributions illustrate the distributions of pairwise sequence differences (Ramos-Onsins and Rozas, 2002). Therefore, the results obtained might be the product of the elimination of less common alleles, and insufficient time to restore genetic equilibrium (Hartl, 2000) within populations. Vector control interventions in Loja might have eliminated a significant number of bugs from populations after spraying infested houses, driving the populations into a bottleneck. The apparent absence of sylvatic populations (Grijalva and Villacis, 2009) that can serve as sources of re-infestation supports this hypothesis.

As mentioned before, vector control interventions were conducted in Loja and Manabí. Insecticide spraying follows entomological searches when infestation has been found (Grijalva and Villacis, 2009). However, in contrast with *P. howardi* in Manabí, no sylvatic bugs have been found in Loja, suggesting that *P. chinai* is a strictly domestic vector. Therefore, vector control interventions may have a great impact on the populations by inducing frequent bottlenecks, which do not facilitate an increase genetic variability. Consequently, re-infestation events found in Loja are most likely due to bugs that survived or avoided insecticide spraying, or are bugs that escaped the intervention in one house given incomplete effects of the insecticide or those that migrated from another
house in the vicinity. However, the possibility of the existence of sylvatic populations of P. chinai should not be eliminated in spite of the lack of such finding.

3.4.2.2 Genetic Structure and geographical dispersal of P. chinai in the Loja province

$F_{ST}$ values were calculated to measure the degree of genetic differentiation among populations of P. chinai in the Loja using Cyt b and ITS-2 loci. Both markers showed no significant genetic structure within this evolutionary unit, suggesting that there is high dispersal among populations of P. chinai. In spite of the lack of significance, it is important to note that $F_{ST}$ values can increase when populations undergo recurrent local extinction and recolonization events (Whitlock and McCuley, 1999). As previously mentioned, triatomines can disperse by flying more than 500 m (WHO, 2002) as well as by walking (Cerere et al., 2004). Most of the populations sampled in the Loja province have a neighbor population within a distance of 10 km, and some others (such as Chaquizca, Guara, Santa Ester and Chirimoyos) within a distance of 2 km.

Therefore, dispersal might be occurring through means of the bug itself or through human assistance. Furthermore, the apparent absence of sylvatic populations suggests that P. chinai in the Loja province has become completely adapted to the human environment, which might facilitate the transmission of T. cruzi in endemic areas. If the parasite is not present among humans or domestic and synanthropic mammals found in the peri-domestic habitats on which P. chinai feeds, the probability of transmission due to
this particular evolutionary unit might decrease. However, the possibility of the presence or future emergence of sylvatic populations should not be discarded.

The shape and distribution of the Cyt b haplotype network shows the wide geographical distribution of a limited number of haplotypes (H = 10) of Cyt b *P. chinai*, which are genetically closely related (Fig 3.8). In addition, the networks constructed using the ITS-2 samples show a similar pattern to Cyt b. However, a wider range of genetic distances is found among these haplotypes (Fig 3.9). Therefore, it is clear that the distribution of haplotypes among the populations lacks genetic structure, which suggests that there is a high level of dispersal of *P. chinai* in Loja, and possible gene flow among populations. For example, Cyt b haplotypes like H2 and H10 (Fig 3.8) are distributed in more than one community, indicating lack of genetic structuring among communities, and possible gene flow. In support of this hypothesis, the ITS-2 haplotypes H13 and H2 (Fig 3.9) are also distributed in more than one community within the province. Bugs found re-infesting these houses possess the same Cyt b haplotypes as the bugs from the previous year infestation. Therefore, it is possible that insecticide spraying is not killing bugs, which can be hidden in places where the insecticide cannot reach (i.e. deep wall cracks, clothes and utensils removed before insecticide spraying). These bugs left behind might then reproduce, but not develop resistance since they were not exposed to the insecticide deltamethrin. In addition, the lack of genetic structure could also be explained by a recent colonization by bugs originated from the same biological source.

The comparisons of the geographic (km) and genetic ($F_{ST}$) distances among populations distributed in Loja were conducted based on the isolation-by-distance model,
the results showed a significant correlation \((p < 0.01)\) between the distances. However the correlation was not strong \((r^2 = 0.244)\). Therefore, because ITS-2 did not show any correlation, it is possible that *P. chinai* is gaining some degree of structure with time, which can be only detected by Cyt b. Such genetic structure might be attributed to genetic drift.

### 3.4.3 Phylogeography of *Panstrongylus rufotuberculatus*

#### 3.4.3.1 Genetic diversity and population demography of *P. rufotuberculatus* in Loja and Manabí.

*Panstrongylus rufotuberculatus* is found in sympatry with *P. howardi* and *P. chinai* in the Manabí and Loja respectively. The samples collected for this study were found in sylvatic, domestic, and peri-domestic habitats of Manabí, but only in domestic habitats in Loja. Contrary to what was observed for *P. howardi* and *P. chinai* in Loja and Manabí, *P. rufotuberculatus* shows notably higher genetic diversity for Cyt b and ITS-2 (Table 3.1). Among the 12 samples collected from both provinces, 11 Cyt b and 6 ITS-2 haplotypes were found. In addition, among 8 samples collected in Manabí, 7 Cyt b and 4 ITS-2 haplotypes were found.

Neutrality tests conducted on samples of the species from both provinces showed that population stability might better describe the demographic status of *P. rufotuberculatus* in the two provinces. In addition, mismatch distributions departed from the expected Poisson distribution under the population expansion model (Fig 3.15). When neutrality test were conducted using samples from the Manabí only, the results showed
the same demographic trend (Table 3.4). The differences in genetic variability and demography between *P. rufotuberculatus* and the other two species might be related to a preference of *P. rufotuberculatus* for sylvatic instead of domestic or per domestic habitats. The presence of a low number of specimens collected in domestic habitats, compared to the number of samples of the other two evolutionary units, suggests such preference. Therefore, it is possible that the genetic variability and demography of populations of *P. rufotuberculatus* are not as affected by vector control as the other two evolutionary units because this species does not frequently infest or colonize domestic habitats and is, acting only an occasional visitor.

### 3.4.3.2 Genetic Structure and geographical dispersal of *P. rufotuberculatus* in the Loja and Manabí provinces

No genetic structure was found among populations of *P. rufotuberculatus*. Even though *F*<sub>ST</sub> values for Cyt b and ITS-2 were close to 1, no statistical significance was found (Table 3.7; Table 3.8). Such lack of significance is most probably the result of comparing populations where the sample sizes were too small (n < 5) to confer power to the test.

The lack of correlation found in the tests of isolation by distance conducted using Cyt b and ITS-2 markers suggests that this species has recently dispersed from one province to the other. Cytochrome b data showed no correlation between genetic and geographic distances among members of this species distributed in both Manabí and Loja. However, a weak correlation (*r*<sup>2</sup> = 0.23) was found when ITS-2 sequences were
compared \((p < 0.001)\). Such results suggest that females have dispersed more than males between the two provinces. Because of the wide distance separating Loja and Manabí (> 300 km), the dispersal of \(P. rufotuberculatus\) could be explained by human assistance. Nevertheless, no current data in regard of the presence of triatomines in the Guayas, El Oro and Los Rios provinces, located between Loja and Manabi have been collected.

Two lineages were recovered when Cyt b haplotypes were used to construct a parsimony network of \(P. rufotuberculatus\) from Loja and Manabi. However, these lineages do not correspond to any specific geographical location in terms of population or province, which supports the hypothesis that \(P. rufotuberculatus\) has recently dispersed between the two provinces. In addition, the ITS-2 haplotype networks constructed revealed haplotypes distributed in both provinces, indicating that dispersal of males is also occurring, but perhaps in a lower degree. Because no sylvatic bugs have been found in Loja, assisted transport of \(P. rufotuberculatus\) from Manabi to Loja can be hypothesized.

\(Panstrongylus rufotuberculatus\) is one of the species of triatomines able to transmit \(T. cruzi\). This species has been found in sylvatic, domestic, and peri-domestic habitats in Manabi, and the small number of bugs found in the domestic and peri-domestic habitats suggest that this species might be an occasional visitor. However, the finding that \(P. rufotuberculatus\) haplotypes are dispersed between Loja and Manabi, and that no sylvatic populations of this species have been found in Loja, suggests that the species is being domesticated in this province. This might represent a serious problem for vector control because these bugs could become new infesting vectors after other species
of triatomines are eliminated by the insecticide spraying. However, given that these populations might be exclusively domestic, insecticide spraying can be more effective in controlling this species than controlling those species living in sylvatic habitats, and are available to infest the domicile.

3.4.4 Implications for the transmission of Chagas disease

Triatomines, including *Panstrongylus sp.* are the vectors that transmit *T. cruzi*, the causative agent of Chagas disease (Lewinsohn, 2003). The information provided by this study can be used to design and implement vector control programs, as well as to monitor and track the possible sources of transmission of *T. cruzi* among rural communities. For example, Ocaña-Mayorga et al., (2010) conducted a population study of *T. cruzi* found in triatomines and small mammals in the same Loja communities I studied and found that there are two lineages (sylvatic and domestic/peri-domestic) of *T. cruzi* strain TcI, distributed among populations in the communities.

The information resulting from my study can be used to associate a specific parasite lineage to a vector species, which will serve to establish priorities in terms of the vector that should be controlled. For example, Ocaña-Mayorga et al. (2010) found that the sylvatic lineage of TcI was found in domestic and peri-domestic habitats of the Santa Ester community, suggesting that the source of the parasite found is sylvatic (triatomine or mammal). My study showed that *P. chinai* can be found in Santa Ester infesting domestic and peri-domestic habitats. However, no sylvatic focus of *P. chinai* has been found so far. Therefore, it is possible that other species of triatomines that have been
found infesting these habitats in this community could be responsible for the transmission of *T. cruzi*. The monitoring of the vector/parasite dynamics should be carefully evaluated taking into account the biology of the triatomine and the means of parasite transmission among these bugs. For example, the parasite can be transmitted through the direct feeding of one triatomine from the body of another one (Eldrige and Endman, 2004), and two different lineages of *T. cruzi* might be detected in one bug.

### 3.4.5 Implications for the control of Chagas Disease

I found that *P. howardi* in Manabi has experienced a recent population expansion, and populations are gaining genetic structure according to the isolation-by-distance model. In addition, this study suggests that sylvatic populations of this evolutionary unit are responsible for re-infestations after insecticide treatment, probably because the bugs in this area are adapting their life cycle to partially sylvatic and partially domestic stages. Triatomines dwelling in sylvatic microhabitats like palm trees and nests in the vicinity of houses might be able to fly towards the peri-domestic habitat in search of a meal, providing the source of re-infestation. After feeding, they could fly back to the sylvatic habitat to deposit their eggs. This hypothesis is in agreement with Suarez-Dávalos et al. (2010) findings which shows the sylvatic presence of early nymph stages only, compared to the presence of later nymph stages and adults only in peri-domiciliary habitats in Manabí province, suggesting the colonization of sylvatic habitats and the migration of late-stage nymphs to peri-domestic areas. This could explain the range expansion
suggested by the present study given that new sylvatic as well as new domestic habitats can be visited frequently by the bugs.

Panstrongylus chinai, in contrast with P. howardi, was found to be an evolutionary unit that has been domesticated in Loja and which has been affected by concurrent vector control interventions. However, this study found evidence suggesting that insecticide spraying might not eliminate all the bugs in an infested house and that insecticide spraying only in infested houses might allow escaping bugs to infest previously uninfested neighboring houses or peri-domestic habitats. In fact, Grijalva MJ (unpublished) observed adult triatomines flying inside the houses attracted by the light, which origin is unknown.

Finally, it was found that P. rufotuberculatus might be an occasional domestic and peri-domestic visitor in Manabí. This species has the potential to become exclusively domestic, as seen in Loja. However, the possible existence of sylvatic populations that have not been recorded cannot be eliminated. Because of the apparently recent dispersal of this species from one province to the other, and the small sample sizes, it is difficult to determine the impact of vector control interventions on P. rufotuberculatus. However, the higher genetic variability of this species compared to the other two evolutionary units I studied suggests that the impact of these interventions is not strong.

Different vector control interventions should be designed to eliminate triatomine species from houses in Loja and Manabí. However, such redesign might imply the increase of the frequency of insecticide spraying or the use of more toxic insecticides. Such interventions, in the end, would not permanently prevent triatomine infestation
because sylvatic populations cannot be eliminated and other species will re-infest houses after the previous species has been controlled, or the bugs might develop resistance to the insecticide used. Therefore, the protection of the residents of the area by the improvement of living conditions, and the construction of better infrastructure that does not allow the entry and colonization of bugs into houses, might be a better approach to control the incidence of Chagas disease. In addition, the establishment of new communities should include the prevention of the destruction of the sylvatic habitats, which sustain the natural life cycle of triatomines.

3.4.6 The context of this study

Previous phylogeographic studies have focused on vectors of Chagas disease in the Triatoma and Rhodnius genera (Monteiro et al., 2003; Bargues et al., 2008). These studies have been conducted using samples collected in the geographical range between Mexico through Argentina. In addition, a number of genetic markers such as ITS-2, D2 (Ribosomal unit), CO1 (Cytochrome Oxidase 1), and microsatellites have been used to infer the phylogeography of these species in different regions of the American continent (Piccinali et al., 2009; Marcet et al., 2006; Monteiro et al., 2003; Bargues et al; 2008). The results of these studies have provided valuable information about the demography and geographic dispersal status of the triatomine populations, which serves as a basis to design vector control programs, targeting a particular species.

In Ecuador, the study of the population genetics and phylogeography of triatomines is relatively recent, although, previous studies have analyzed the
biogeography of a number of species (Abad-Franch et al., 2001). The microdistribution of *P. howardi* and *P. rufotuberculatus* in sylvatic habitats in the Manabí province has been recently studied (Suarez-Dávalos et al., 2010). These results are complemented by the results of the present phylogeographic study of the species, which describe the demographic events that may be shaping the genetic structure of this species and evolutionary unit in the Manabí province. Both studies agree in the hypothesis that sylvatic habitats are the source of reinestation of houses by *P. howardi*, and that the design of the current vector control strategy might not be effective in eliminating the bugs in this region. In addition, reports indicating the sylvatic presence of *P. rufotuberculatus* in Manabí (Suarez-Dávalos et al., 2010), and their absence in the same habitats in Loja (Grijalva and Villacis, 2009) could suggest that the bugs found in the different regions belong to different genetic lineages. However, the results of this study bring about a different hypothesis, suggesting the recent passive transport of *P. rufotuberculatus* from one province to the other, given the genetic similarity and wide dispersal of this species between the two regions. Similarly, triatomine surveys conducted in Loja have also found the apparent absence of *P. chinai* from sylvatic habitats of this province (Grijalva and Villacis, 2009). The information provided by the phylogeographic study of this evolutionary unit suggests that vector control might be more effective in eliminating these bugs in this region.

In conclusion, it is clear that the in-depth study of the populations’ dynamics of triatomines in different geographic regions provides valuable information that can benefit vector control. However, this information should be accompanied by the study of the
biology of the different species because, as previously noted, the biology of the bugs can vary according to the species or evolutionary unit and the geographical region where they are found. The combination of such studies is of vital importance to design or continue effective vector control programs to prevent the transmission of *T. cruzi* and decrease the incidence of Chagas disease in Ecuador.
REFERENCES

http://www.cdc.gov/parasites/chagas/biology.html


http://www.geoportaligm.gob.ec/index2.html


*Prospectus in Biology and Medicine.* 46(4):532-549.


LNS (Ed) 2009 Atlas LNS 3ra edition Don Bosco pg 45.


Ngan E. 2006. Isolation by Distance Web service with incorporation of Data sets. Thesis presented to the Faculty of San Diego State University.


Rambaut A. 2009. FigTree Tree Figure Drawing tool. Institute of Evolutionary Biology. University of Edinburgh.


Table 3.1

Genetic indices calculated among *Panstrongylus* sp. at the intra-specific level using cytochrome b and ITS-2 sequences

<table>
<thead>
<tr>
<th>Species/Regions</th>
<th>No. pop.</th>
<th>No. sam.</th>
<th>Cytochrome b</th>
<th>ITS-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$S^a$</td>
<td>$H^b$</td>
</tr>
<tr>
<td><em>P. howardi</em></td>
<td>8</td>
<td>33</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td><em>P. chinai</em></td>
<td>12</td>
<td>26</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em>, Manabí</td>
<td>6</td>
<td>8</td>
<td>57</td>
<td>7</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em>, Loja and Manabí</td>
<td>9</td>
<td>12</td>
<td>59</td>
<td>11</td>
</tr>
</tbody>
</table>

$^a$ $S$ = number of segregating sites in a sample

$^b$ $H$ = number of haplotypes found in a sample

$^c$ $h$ = haplotype diversity (the probability that two randomly sampled sequences are different)

$^d$ $k$ = sequence diversity (the average number of nucleotide differences between pairs of sequences)

$^e$ $\pi$ = nucleotide diversity (the probability that two randomly sampled homologous nucleotides are different)
Table 3.2

$F_{ST}$ values calculated using Cyt b sequences of *P. howardi* samples collected in 12 rural communities of Manabí province.

<table>
<thead>
<tr>
<th></th>
<th>El Bejuco</th>
<th>La Cienega</th>
<th>La Encantada</th>
<th>Maconta Abajo</th>
<th>San Gabriel</th>
<th>Zapallo</th>
<th>Santa Rosa de las Palmas</th>
<th>Pacoche</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Bejuco</td>
<td></td>
<td>0.68359</td>
<td>0.99902</td>
<td>0.54004</td>
<td>0.23438</td>
<td>0.33594</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>La Cienega</td>
<td>-0.018</td>
<td></td>
<td></td>
<td>0.17773</td>
<td>0.01855</td>
<td>0.00586</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>La Encantada</td>
<td>-0.174</td>
<td></td>
<td></td>
<td>0.34961</td>
<td>0.09961</td>
<td>0.46875</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Maconta Abajo</td>
<td>-0.144</td>
<td>0.62791</td>
<td>0.25</td>
<td>0.31055</td>
<td>0.55176</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>San Gabriel</td>
<td>0.191</td>
<td>1</td>
<td>1</td>
<td>0.66667</td>
<td>0.08301</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Zapallo</td>
<td>0.032</td>
<td>0.25466</td>
<td>0.00826</td>
<td>-0.02515</td>
<td>0.28571</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Santa Rosa de las Palmas</td>
<td>-0.969</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-0.71429</td>
<td>0.99902</td>
<td></td>
</tr>
<tr>
<td>Pacoche</td>
<td>0.749</td>
<td>1</td>
<td>1</td>
<td>0.89474</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Values below diagonal = $F_{ST}$ values calculated using pair wise differences.
Values above diagonal = $p$-values
Shaded values below diagonal = significant $F_{ST}$ values
Shaded values above diagonal = corresponding $p$-value for significant $F_{ST}$
Table 3.3

$F_{ST}$ values calculated using ITS-2 sequences of *P. howardi* samples collected in 8 rural communities of Manabí province.

<table>
<thead>
<tr>
<th></th>
<th>El Bejuco</th>
<th>La Cienega</th>
<th>La Extensa</th>
<th>Maconta Abajo</th>
<th>San Gabriel</th>
<th>Zapallo</th>
<th>Pacoche</th>
<th>Santa Rosa de las Palmas</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Bejuco</td>
<td></td>
<td>0.60156</td>
<td>0.44238</td>
<td>0.45996</td>
<td>0.02734</td>
<td>0.46875</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>La Cienega</td>
<td>-0.03209</td>
<td></td>
<td>0.31738</td>
<td>0.66895</td>
<td>0.22852</td>
<td>0.33984</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>La Extensa</td>
<td>-0.04987</td>
<td>-0.03325</td>
<td></td>
<td>0.11426</td>
<td>0.11426</td>
<td>0.38477</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Maconta Abajo</td>
<td>-0.13615</td>
<td>-0.16739</td>
<td>0.49231</td>
<td></td>
<td>0.34473</td>
<td>0.45410</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>San Gabriel</td>
<td>0.63107</td>
<td>0.24306</td>
<td>0.97734</td>
<td>0.94631</td>
<td></td>
<td>0.04980</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Zapallo</td>
<td>-0.01566</td>
<td>0.03016</td>
<td>0.02174</td>
<td>-0.06557</td>
<td>0.91146</td>
<td></td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Pacoche</td>
<td>-0.54516</td>
<td>-0.65451</td>
<td>0.6</td>
<td>-0.16667</td>
<td>0.98675</td>
<td>-0.29032</td>
<td></td>
<td>0.99902</td>
</tr>
<tr>
<td>Santa Rosa de las Palmas</td>
<td>-0.72096</td>
<td>-0.70163</td>
<td>-1</td>
<td>-0.07692</td>
<td>0.98621</td>
<td>-0.73913</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Values below diagonal = $F_{ST}$ values calculated using pair wise differences.
Values above diagonal = $p$-values
Table 3.4

Neutrality tests and demographic history of *Panstrongylus sp.* from Loja and Manabí provinces using Cyt b and ITS-2 markers.

<table>
<thead>
<tr>
<th>Species/Regions</th>
<th>No. pop</th>
<th>No. sam.</th>
<th>Cytochrome b</th>
<th>ITS-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>R²f</td>
<td>Fs</td>
</tr>
<tr>
<td><em>P. howardi</em></td>
<td>8</td>
<td>33</td>
<td>0.0496&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-5.483&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>P. chinai</em></td>
<td>12</td>
<td>26</td>
<td>0.0974</td>
<td>-2.189</td>
</tr>
<tr>
<td><em>P. rufotuberculatus, Manabi</em></td>
<td>6</td>
<td>8</td>
<td>0.2352</td>
<td>1.423</td>
</tr>
<tr>
<td><em>P. rufotuberculatus Loja and Manabi</em></td>
<td>9</td>
<td>12</td>
<td>0.2165</td>
<td>-0.378</td>
</tr>
</tbody>
</table>

<sup>a, b</sup> p < 0.005  
<sup>c</sup> p < 0.05  
<sup>d</sup> p < 0.0001  
<sup>e</sup> p < 0.00001  
<sup>f</sup> R² = Ramos-Onsins and Rozas (2002). Null hypothesis = population stability. Significant results are indicative of demographic expansion.  
<sup>g</sup> Fs = Fu’s F (Fu, 1997) Null hypothesis = population stability. Significant results are indicative of demographic expansion.
h MMSSD = Sum of Squared Deviation between observed and expected mismatch distributions under a sudden demographic expansion model. Significant values are indicative of population stability.

Table 3.5

$F_{ST}$ values calculated using Cyt b sequences of $P. chinai$ samples collected in 12 rural communities of Loja province.

<table>
<thead>
<tr>
<th></th>
<th>Ashimino</th>
<th>Bramadero</th>
<th>Coamine</th>
<th>Chaquizca</th>
<th>Chirimoyos</th>
<th>La Extensa</th>
<th>Guara</th>
<th>Jaguay</th>
<th>Lucarqui</th>
<th>Naranjo Dulce</th>
<th>San Francisco</th>
<th>Santa Ester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashimino</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.08203</td>
<td>0.10352</td>
<td>0.10449</td>
<td>0.13281</td>
<td>0.19238</td>
<td>0.99902</td>
<td>0.73926</td>
<td>0.99902</td>
<td>0.11719</td>
<td></td>
</tr>
<tr>
<td>Bramadero</td>
<td>0</td>
<td>0.99902</td>
<td>0.10156</td>
<td>0.08594</td>
<td>0.09863</td>
<td>0.11133</td>
<td>0.09473</td>
<td>0.99902</td>
<td>0.37695</td>
<td>0.99902</td>
<td>0.12207</td>
<td></td>
</tr>
<tr>
<td>Coamine</td>
<td>0.818</td>
<td>0.99902</td>
<td>0.22852</td>
<td>0.24512</td>
<td>0.75488</td>
<td>0.30566</td>
<td>0.30762</td>
<td>0.99902</td>
<td>0.30273</td>
<td>0.99902</td>
<td>0.34277</td>
<td></td>
</tr>
<tr>
<td>Chaquizca</td>
<td>0.8</td>
<td>0.90909</td>
<td>0.6</td>
<td>0.41406</td>
<td>0.42480</td>
<td>0.36230</td>
<td>0.07129</td>
<td>0.99902</td>
<td>0.08887</td>
<td>0.99902</td>
<td>0.42480</td>
<td></td>
</tr>
<tr>
<td>Chirimoyos</td>
<td>0.875</td>
<td>1</td>
<td>0.5</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.10938</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.09180</td>
<td>0.99902</td>
<td>0.42480</td>
<td></td>
</tr>
<tr>
<td>La Extensa</td>
<td>0.636</td>
<td>0.75</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.99902</td>
<td>0.09375</td>
<td>0.99902</td>
<td>0.09668</td>
<td>0.99902</td>
<td>0.99902</td>
<td></td>
</tr>
<tr>
<td>Guara</td>
<td>0.836</td>
<td>1</td>
<td>1</td>
<td>0.36842</td>
<td>0.0</td>
<td>0.31934</td>
<td>0.99902</td>
<td>0.33105</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td></td>
</tr>
<tr>
<td>Jaguay</td>
<td>0.739</td>
<td>1</td>
<td>1</td>
<td>0.88</td>
<td>1</td>
<td>0.67568</td>
<td>1</td>
<td>0.99902</td>
<td>0.31641</td>
<td>0.99902</td>
<td>0.31738</td>
<td></td>
</tr>
<tr>
<td>Lucarqui</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.33333</td>
<td>1</td>
<td>1</td>
<td>0.69043</td>
<td>0.99902</td>
<td>0.32715</td>
<td></td>
</tr>
<tr>
<td>Naranjo Dulce</td>
<td>0.045</td>
<td>0.25</td>
<td>0.77778</td>
<td>0.8125</td>
<td>0.90323</td>
<td>0.63415</td>
<td>0.8</td>
<td>0.33333</td>
<td>0.99902</td>
<td>0.35645</td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Francisco</td>
<td>0.882</td>
<td>1</td>
<td>1</td>
<td>0.81818</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0.84615</td>
<td></td>
<td>0.37793</td>
<td></td>
</tr>
<tr>
<td>Santa Ester</td>
<td>0.836</td>
<td>1</td>
<td>1</td>
<td>0.36842</td>
<td>0</td>
<td>-0.2</td>
<td>0</td>
<td>1</td>
<td>0.85714</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values below diagonal = $F_{ST}$ values calculated using pair wise differences.
Values above diagonal = $p$-values
Table 3.6

$F_{ST}$ values calculated using ITS-2 sequences of *P. chinai* samples collected in 12 rural communities of Loja province.

<table>
<thead>
<tr>
<th></th>
<th>Ashimingo</th>
<th>Bramaderos</th>
<th>Coamine</th>
<th>Chaquizca</th>
<th>Chirimoyos</th>
<th>La Extensa</th>
<th>Guara</th>
<th>Jaguay</th>
<th>Lucarqui</th>
<th>Naranjo Dulce</th>
<th>San Francisco</th>
<th>Santa Ester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashimingo</td>
<td>0.20410</td>
<td>0.99902</td>
<td>0.10156</td>
<td>0.38770</td>
<td>0.10059</td>
<td>0.61914</td>
<td>0.22852</td>
<td>0.99902</td>
<td>0.61230</td>
<td>0.99902</td>
<td>0.20801</td>
<td></td>
</tr>
<tr>
<td>Bramaderos</td>
<td>0.19896</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.10547</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.09473</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.62305</td>
<td></td>
</tr>
<tr>
<td>Coamine</td>
<td>-0.24099</td>
<td>0.96957</td>
<td>0.24805</td>
<td>0.99902</td>
<td>0.28223</td>
<td>0.99902</td>
<td>0.39844</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.36133</td>
<td></td>
</tr>
<tr>
<td>Chaquizca</td>
<td>0.20429</td>
<td>0</td>
<td>0.98678</td>
<td>0.42676</td>
<td>0.99902</td>
<td>0.31738</td>
<td>0.90723</td>
<td>0.99902</td>
<td>0.11426</td>
<td>0.99902</td>
<td>0.77832</td>
<td></td>
</tr>
<tr>
<td>Chirimoyos</td>
<td>-0.03427</td>
<td>0.4871</td>
<td>-0.98693</td>
<td>0.49013</td>
<td>0.38806</td>
<td>0.75781</td>
<td>0.19531</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.40234</td>
<td></td>
</tr>
<tr>
<td>La Extensa</td>
<td>0.20446</td>
<td>0</td>
<td>0.98242</td>
<td>-0.10526</td>
<td>0.48852</td>
<td>0.41504</td>
<td>0.91309</td>
<td>0.99902</td>
<td>0.08691</td>
<td>0.99902</td>
<td>0.68555</td>
<td></td>
</tr>
<tr>
<td>Guara</td>
<td>-0.29259</td>
<td>0.23879</td>
<td>-0.98693</td>
<td>0.24088</td>
<td>-0.6008</td>
<td>0.2385</td>
<td>0.35352</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td></td>
</tr>
<tr>
<td>Jaguay</td>
<td>0.03056</td>
<td>-0.27397</td>
<td>0.9934</td>
<td>-0.07143</td>
<td>0.36596</td>
<td>-0.09615</td>
<td>0.00649</td>
<td>0.99902</td>
<td>0.32520</td>
<td>0.99902</td>
<td>0.99902</td>
<td></td>
</tr>
<tr>
<td>Lucarqui</td>
<td>-0.23543</td>
<td>0.96957</td>
<td>1</td>
<td>0.98675</td>
<td>-0.96129</td>
<td>0.98242</td>
<td>-0.97403</td>
<td>0.9934</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.37500</td>
<td></td>
</tr>
<tr>
<td>Naranjo Dulce</td>
<td>-0.28477</td>
<td>0.25465</td>
<td>-1</td>
<td>0.26073</td>
<td>-0.59509</td>
<td>0.2583</td>
<td>-0.95484</td>
<td>0.02564</td>
<td>0.98684</td>
<td>0.99902</td>
<td>0.33496</td>
<td></td>
</tr>
<tr>
<td>San Francisco</td>
<td>0.06926</td>
<td>0.95732</td>
<td>1</td>
<td>0.98148</td>
<td>0.45422</td>
<td>0.97523</td>
<td>0.08709</td>
<td>0.9907</td>
<td>1</td>
<td>0.08761</td>
<td>0.29004</td>
<td></td>
</tr>
<tr>
<td>Santa Ester</td>
<td>0.02651</td>
<td>-0.01695</td>
<td>0.97351</td>
<td>-0.2</td>
<td>0.35461</td>
<td>-0.06329</td>
<td>-0.00971</td>
<td>0</td>
<td>0.97342</td>
<td>0.01274</td>
<td>0.96313</td>
<td></td>
</tr>
</tbody>
</table>

Values below diagonal = $F_{ST}$ values calculated using pair wise differences.
Values above diagonal = *p*-values
Table 3.7

$F_{ST}$ values calculated using Cyt b sequences of *P. rufotuberculatus* samples collected in 9 rural communities of Loja and Manabí provinces.

<table>
<thead>
<tr>
<th></th>
<th>Algarobillo, L</th>
<th>Pitayo, L</th>
<th>El Sauce, L</th>
<th>El Bejuco, M</th>
<th>San Francisco, M</th>
<th>Estero Seco, M</th>
<th>Humedad, M</th>
<th>Punta Larga, M</th>
<th>San Ramon, M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algarobillo, L</td>
<td></td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.33301</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Pitayo, L</td>
<td>-0.725</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>El Sauce, L</td>
<td>-0.76</td>
<td>1</td>
<td></td>
<td>0.99902</td>
<td>0.66504</td>
<td>0.64551</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>El Bejuco, M</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>0.99902</td>
<td>0.32715</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>San Francisco, M</td>
<td>-0.936</td>
<td>-0.88</td>
<td>-0.77358</td>
<td>-1</td>
<td></td>
<td>0.66016</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Estero Seco, M</td>
<td>0.056</td>
<td>0.86139</td>
<td>0.17647</td>
<td>0.66</td>
<td>0.01818</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humedad, M</td>
<td>-0.833</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0.85567</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Punta Larga, M</td>
<td>-0.833</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0.85859</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Ramon, M</td>
<td>-0.66</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.80769</td>
<td>0.86408</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Values below diagonal = $F_{ST}$ values calculated using pair wise differences.
Values above diagonal = $p$-values
L = Loja
M = Manabí
Table 3.8

$F_{ST}$ values calculated using ITS-2 sequences of *P. rufotuberculatus* samples collected in 9 rural communities of Loja and Manabí provinces.

<table>
<thead>
<tr>
<th></th>
<th>Algarobillo, L</th>
<th>Pitayo, L</th>
<th>El Sauce, L</th>
<th>El Bejuco, M</th>
<th>San Francisco, M</th>
<th>San Ramon, M</th>
<th>Estero Seco, M</th>
<th>Humedad, M</th>
<th>Punta Larga, M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algarobillo, L</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.32324</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Pitayo, L</td>
<td>-0.98198</td>
<td></td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>El Sauce, L</td>
<td>0.41333</td>
<td>1</td>
<td>1</td>
<td>0.99902</td>
<td>0.28320</td>
<td>0.99902</td>
<td>0.30664</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>El Bejuco, M</td>
<td>-0.89655</td>
<td>1</td>
<td>1</td>
<td>0.33008</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.69629</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>San Francisco, M</td>
<td>0.00901</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>San Ramon, M</td>
<td>0.36416</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.30762</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Estero Seco, M</td>
<td>0.03084</td>
<td>-1</td>
<td>0.96882</td>
<td>-0.27273</td>
<td>0</td>
<td>0.94636</td>
<td>-1</td>
<td>0.99902</td>
<td>0.99902</td>
</tr>
<tr>
<td>Humedad, M</td>
<td>-0.91304</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Punta Larga, M</td>
<td>-0.91304</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Values below diagonal = $F_{ST}$ values calculated using pair wise differences.
Values above diagonal = $p$-values
L = Loja
M = Manabí
Figure 1.1. Species of Panstrongylus included in the present study. A. Panstrongylus howardi. B. P. rufotuberculatus. C. P. chinai.
Figure 2.1. Map of Ecuador. Topology is indicated as gradient gray scale lines according to the altitude they represent. Heavy borders represent the borders of Loja and Manabí provinces, included in this study.
Figure 2.2. Representative gel (1% agarose) showing the PCR products obtained from the amplification of Cytochrome b DNA (primers Cytb7432 - Cytb7433). *Panstrongylus* cytochrome b DNA is amplified as a typical band of 685 bp. No additional bands are expected. MM= molecular marker; NC= negative control; Lanes 1 – 4 = samples from *Panstrongylus sp.*
Figure 2.3. Representative gel (1% agarose) showing the PCR products obtained from the amplification of the rDNA cistron located between the 5.8S and 28S rDNA (primers F5.8S and R28S), which contains the ITS-2 DNA fragment. Given the repetitive nature of this fragment, Panstrongylus rDNA is amplified in different sized bands. MM= molecular marker; NC= negative control; Lanes 1 – 6 = samples from Panstrongylus sp.
Figure 2.4. Representative gel (1% agarose) showing the products obtained from the alkaline plasmid purification and EcoRI digestion of the rDNA cistron cloned into a TOPOTA cloning vector. Given the repetitive nature of this fragment, *Panstrongylus* rDNA is amplified in different sized bands. Plasmids containing the largest fragments were selected for a second purification; MM10 = 10 kb DNA ladder; Lanes 1 – 3 = EcoRI digested samples from *Panstrongylus sp.*; Lanes 4 – 6 = undigested plasmid.
Figure 2.5. Representative gel (1% agarose) showing the products obtained from the plasmid purification using Qiagen Miniprep Kit and EcoRI digestion of the rDNA cistron cloned into a TOPOTA cloning vector. Given the repetitive nature of this fragment, *Panstrongylus* rDNA is amplified in different sized bands. Longest fragments were used for DNA sequencing. MM10 = 10 kb molecular marker; MM1= 1kb molecular marker; Lanes 1 – 5 =clones obtained from two different samples.
Figure 2.6. Plot of points representing convergence of data and trees used for inferring phylogeny. A. Cyt b. B. ITS-2. C. Concatenated Cyt b and ITS-2 data. D. ITS-2 South American Samples of *Panstrongylus*. 
Figure 2.7. Phylogram tree of *Panstrongylus* sp. found in Loja and Manabi provinces using Cytochrome b sequences. Values below branches are posterior probabilities from the Bayesian analysis using MrBayes. Values above branches are likelihood values obtained after ML analysis using GARLI. Blue branches represent samples collected in Loja province. Black branches represent samples collected in Manabi province. Scale bar indicates the number of nucleotide substitutions per sequence position.
Figure 2.8. Phylogram of *Panstrongylus* sp. found in Loja and Manabi provinces using ITS-2 sequences. Values below branches are posterior probabilities from the Bayesian analysis using MrBayes. Values above branches are likelihood values obtained after ML analysis using GARLI. Blue branches represent samples collected in Loja province. Black branches represent samples collected in Manabi province. Scale bar indicates the number of nucleotide substitutions per sequence position.
Figure 2.9. Phylogram of *Panstrongylus* sp. found in Loja and Manabí provinces using concatenated data of Cytochrome b and ITS-2 sequences. Blue branches represent samples collected in Loja province. Black branches represent samples collected in Manabí province. Values above branches are posterior probabilities from the Bayesian analysis using MrBayes. Scale bar indicates the number of nucleotide substitutions per sequence position.
Figure 2.10. Phylogram inferred from *Panstrongylus sp.* found in South America using ITS-2 sequences. Dashed lines and support value in blue represent branches supported by ML method. Values below branches are posterior probabilities from the Bayesian analysis using MrBayes. Values above branches represent likelihood obtained using ML method. Scale bar indicates the number of nucleotide substitutions per sequence position.
Figure 3.1. Map of the Manabi province. Topologic map of the province indicating its ecoregions (See legend). White circles with black center indicate the location of the populations studied. The names of the communities where bug populations were found are indicated next to each circle. Upper left box indicates the location of the Manabi province in Ecuador.
Figure 3.2. Geographic distribution of *P. howardi* Cyt b haplotype network. *P. howardi* Cyt b haplotype relationship and distribution in the Manabí province. A) Haplotype network constructed using Cyt b sequences of *P. howardi* from the Manabí province. Square represents the haplotype that is considered as the ancestor of the other haplotypes. Circles represent haplotypes found in different localities of the Manabí province. Lines connecting haplotypes, separated by a dot represent the number of mutations in which those haplotypes differ. H1, H4, H9, H12 = El Bejuco; H2, H3, H8 = Zapallo; H5 = La Cienega; H6 = San Gabriel; H7 = Maconta Abajo; H10 = La Cienega, El Bejuco, Zapallo, Santa Rosa de las Palmas, La Encantada, Maconta Abajo; H11 = Pacoche. B) Cyt b haplotype distribution in Manabí province. Circles represent haplotypes found in rural communities in the Manabí province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community.
Figure 3.3. Geographic distribution of *P. howardi* ITS-2 haplotype network. *P. howardi* ITS-2 haplotype relationship and distribution in the Manabi province. Haplotype network constructed using ITS-2 sequences of *P. howardi* from the Manabi province. Square represents the haplotype that is considered as the ancestor of the other haplotypes. Circles represent haplotypes found in different localities of the Manabi province. Lines connecting haplotypes, separated by a dot represent the number of mutations in which those haplotypes differ. H1, H4, H6, H15, H22 = La Cienega; H2, H3, H7, H13, H17 = Zapallo; H5 = Pacoche; H8, H9, H11, H14, H21 = El Bejuco; H10, H12 = Maconta Abajo; H16, H18 = La Encantada; H19 = Santa Rosa de las Palmas, La Encantada, El Bejuco; H20 = La Cienega, El Bejuco; H24 = El Bejuco, San Gabriel, La Cienega. B) ITS-2 haplotype distribution in the Manabi province. Circles represent haplotypes found in rural communities in the Manabi province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community.
Figure 3.4. *Panstrongylus howardi* mismatch distributions. Red dashed line represents the observed mismatch distribution. Green solid line represents the expected mismatch distribution for a population expansion model. A) Cyt b *P. howardi* B) ITS-2 *P. howardi*. 
Figure 3.5. Isolation-by-distance of *P. howardi* using Cyt b sequences. A) IBD comparing *P. howardi* samples collected in Manabí province. B) IBD comparing *P. howardi* samples collected in communities La Cienega, Maconta Abajo, Zapallo and Pacoche. C) IBD comparing *P. howardi* samples collected in El Bejuco, San Gabriel, Zapallo and La Encantada.
Figure 3.6. Isolation-by-distance of *P. howardi* using ITS-2 sequences. A) IBD comparing *P. howardi* samples collected in the Manabí province. B) IBD comparing *P. howardi* samples collected in La Cienega, Maconta Abajo, Zapallo and Pacoche.
Figure 3.7. Map of the Loja province. Topologic map of the province indicating its ecoregions (See legend). White circles with black center indicate the location of the populations studied. The names of the communities where bug populations were found are indicated next to each circle. Lower left box indicates the location of the Loja province in Ecuador.
Figure 3.8. Geographic distribution of *P. chinai* Cyt b haplotype network. *P. chinai* Cyt b haplotype relationship and distribution in the Loja province. A) Haplotype network constructed using Cyt b sequences of *P. chinai* from the Loja province. Square represents the haplotype that is considered as the ancestor of the other haplotypes. Circles represent haplotypes found in different localities of the Loja province. Lines connecting haplotypes, separated by a dot represent the number of mutations in which those haplotypes differ. H1 = San Francisco; H2 = Ashimingo, Bramaderos, Naranjo Dulce; H3 = Naranjo Dulce; H4 = Jaguay; H5 = Ashimingo, Lucarqui; H6, H9 = La Extensa; H7 = Chaquizca; H8 = Coamine; H10 = Guara, Chaquizca, Chirimoyos, La Extensa, Santa Ester. B) Cyt b haplotype distribution in the Loja province. Circles represent haplotypes found in rural communities in the Loja province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community.
Figure 3.9. Geographic distribution of *P. chinai* ITS-2 haplotype networks. *P. chinai* ITS-2 haplotype relationship and distribution in the Loja province. A) Haplotype network constructed using ITS-2 sequences of *P. chinai* from the Loja province. Square represents the haplotype that is considered as the ancestor of the other haplotypes. Circles represent haplotypes found in different localities of the Loja province. Lines connecting haplotypes, separated by a dot represent the number of mutations in which those haplotypes differ. H1 = Jaguay, Bramaderos; H2 = La Extensa, Bramaderos, Jaguay, Chaquizca; H3, H6, H11 = Ashimingo; H4 = Naranjo Dulce; H5 = Bramaderos; H7 = Chirimoyos; H8 = Lucarqui; H9 = Chirimoyos, Naranjo Dulce, Coamine; H10 = Guara; H12, H16 = La Extensa; H13 = Guara, Chirimoyos, Chaquizca, Santa Ester; H14 = Santa Ester; H15 = Chaquizca; H17 = San Francisco. B) ITS-2 haplotype distribution in the Loja province. Circles represent haplotypes found in rural communities in Loja province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community.
Figure 3.10. *Panstrongylus chinai* mismatch distributions. Red dashed line represents the observed mismatch distribution. Green solid line represents the expected mismatch distribution for a population expansion model. A) Cyt b *P. chinai* B) ITS-2 *P. chinai*. 
Figure 3.11. Isolation-by-distance of *P. chinai* using Cyt b sequences. A) IBD comparing *P. chinai* samples collected in Loja province. B) IBD comparing *P. chinai* samples collected in La Extensa, Naranjo Dulce, Ashimingo, Bramaderos, and Jaguay in the Loja province. C) IBD comparing *P. chinai* samples collected in communities Santa Ester, Chirimoyos, Chaquizca and Guara.
Figure 3.12. Isolation-by-distance of *P. chinai* using ITS-2 sequences. A) IBD comparing *P. chinai* samples collected in Loja province. B) IBD comparing *P. chinai* samples collected in communities La Extensa, Naranjo Dulce, Ashimingo, Bramaderos, and Jaguay (Santa Ester*) in Loja province.
Figure 3.13. Geographic distribution of *P. rufotuberculatus* Cyt b haplotype networks. *P. rufotuberculatus* Cyt b haplotype relationship and distribution in Loja and Manabi provinces. A) Cyt b haplotype distribution in Loja province. Circles represent haplotypes found in rural communities in Loja province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community. B) Cyt b haplotype distribution in Manabi province. Circles represent haplotypes found in rural communities in Loja province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community. C) Haplotype network constructed using Cyt b sequences of *P. rufotuberculatus* from Loja and Manabi provinces. Square represents the haplotype that is considered as the ancestor of the other haplotypes. Circles represent haplotypes found in different localities of Loja province. Lines connecting haplotypes, separated by a dot represent the number of mutations in which those haplotypes differ. H1 = El Bejuco, Manabi; H2, H4 = Estero Seco, Manabi; H3 = El Sauce, Loja; H5, H7 = Algarobillo, Loja; H6 = San Francisco, Manabi; H8 = San Ramón, Manabi; H9 = Pitayo, Loja; H10 = Humedad, Manabi; H11 = Punta Larga, San Francisco, Manabi
Figure 3.14. Geographic distribution of *P. rufotuberculatus* ITS-2 haplotype network *P. rufotuberculatus* ITS-2 haplotype relationship and distribution in the Loja and Manabí provinces. A) ITS-2 haplotype distribution in Loja province. Circles represent haplotypes found in rural communities in Loja province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community. B) ITS-2 haplotype distribution in the Manabí province. Circles represent haplotypes found in rural communities in the Loja province. Haplotypes can be connected to more than one community or more than one haplotype can be found in one single community. C) Haplotype network constructed using ITS-2 sequences of *P. rufotuberculatus* from the Loja and Manabí provinces. Square represents the haplotype that is considered as the ancestor of the other haplotypes. Circles represent haplotypes found in different localities of Loja province. Lines connecting haplotypes, separated by a dot represent the number of mutations in which those haplotypes differ. H1, H6 = Algarobillo, Loja; H2 = El Sauce, Loja; H3 = San Ramón, Manabi; H4 = El Bejuco, Manabi; H5 = Humedad, Punta Larga, Estero Seco, Manabi; H7 = San Francisco, Pitayo, Loja; H8 = Estero Seco, Manabi.
Figure 3.15. *Panstrongylus rufotuberculatus* Cyt b mismatch distributions. Red dashed line represents the observed mismatch distribution. Green solid line represents the expected mismatch distribution for a population expansion model. A) *P. rufotuberculatus* Loja and Manabí B) *P. rufotuberculatus* Manabí.
Figure 3.16. Panstrongylus rufotuberculatus ITS-2 mismatch distributions. Red dashed line represents the observed mismatch distribution. Green solid line represents the expected mismatch distribution for a population expansion model. A) *P. rufotuberculatus* Loja and Manabí  B) *P. rufotuberculatus* Manabí.
Figure 3.17. Isolation-by-distance of *P. rufotuberculatus* using Cyt b sequences. A) IBD comparing *P. rufotuberculatus* samples collected in Loja and Manabi provinces. B) IBD comparing *P. rufotuberculatus* samples collected in communities Estero Seco, San Ramon and San Francisco in Manabi province.
Figure 3.18. Isolation-by-distance of *P. rufotuberculatus* using ITS-2 sequences. A) IBD comparing *P. rufotuberculatus* samples collected in Loja and Manabí provinces. B) IBD comparing *P. rufotuberculatus* samples collected in communities Estero Seco, San Ramon and San Francisco in Manabí province.
APPENDICES

Appendix 2.1: List of Triatomines collected in Manabi and Loja provinces in Ecuador and used for the analyses in this study.

<table>
<thead>
<tr>
<th>Species/Trees code</th>
<th>CIEI code</th>
<th>Community</th>
<th>Province</th>
<th>Habitat</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. rufotuberculatus 1</td>
<td>AB4438</td>
<td>Algarobillo</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td>P. rufotuberculatus 2</td>
<td>AB2503</td>
<td>Algarobillo</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td>P. chinai 1</td>
<td>AH2421</td>
<td>Ashimingo</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td>P. chinai 2</td>
<td>AH2408</td>
<td>Ashimingo</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td>P. chinai 3</td>
<td>AH2445</td>
<td>Ashimingo</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td>P. howardi 1</td>
<td>TBJ1528</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 2</td>
<td>BJ5612</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 3</td>
<td>BJ2771</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 4</td>
<td>BJ2997</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Domiciliari</td>
</tr>
<tr>
<td>P. howardi 5</td>
<td>BJ2998</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Domiciliari</td>
</tr>
<tr>
<td>P. howardi 6</td>
<td>BJ2737</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 7</td>
<td>BJ1691</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 8</td>
<td>TBJ1621</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 9</td>
<td>BJ2745</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 10</td>
<td>TBJ1608</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Domiciliari</td>
</tr>
<tr>
<td>P. howardi 11</td>
<td>TBJ1558</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Sylvatic</td>
</tr>
<tr>
<td>P. rufotuberculatus 3</td>
<td>TBJ2238</td>
<td>El Bejuco</td>
<td>Manabí</td>
<td>Domiciliari</td>
</tr>
<tr>
<td>P. chinai 4</td>
<td>BR2260</td>
<td>Bramadero</td>
<td>Loja</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. chinai 5</td>
<td>BR2253</td>
<td>Bramadero</td>
<td>Loja</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. chinai 6</td>
<td>BR2254</td>
<td>Bramadero</td>
<td>Loja</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. chinai 7</td>
<td>CE4586</td>
<td>Coamine</td>
<td>Loja</td>
<td>Domiciliari</td>
</tr>
<tr>
<td>P. howardi 12</td>
<td>CN1349</td>
<td>La Cienega</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 13</td>
<td>CN1350</td>
<td>La Cienega</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 14</td>
<td>CN3519</td>
<td>La Cienega</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 15</td>
<td>CN1360</td>
<td>La Cienega</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>P. howardi 16</td>
<td>CN1352</td>
<td>La Cienega</td>
<td>Manabí</td>
<td>Peri-domiciliari</td>
</tr>
<tr>
<td>Species</td>
<td>Code</td>
<td>Location</td>
<td>Parish</td>
<td>Type</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>--------------</td>
<td>--------</td>
<td>-----------------</td>
</tr>
<tr>
<td><em>P. howardi</em> 17</td>
<td>CN3558</td>
<td>La Cienega</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 18</td>
<td>CN1725</td>
<td>La Cienega</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 19</td>
<td>CN1351</td>
<td>La Cienega</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 8</td>
<td>CQ4349</td>
<td>Chaquizca</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 9</td>
<td>CQ4426</td>
<td>Chaquizca</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 10</td>
<td>CQ4350</td>
<td>Chaquizca</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 11</td>
<td>CY3007</td>
<td>Chirimoyos</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 12</td>
<td>CY3087</td>
<td>Chirimoyos</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 13</td>
<td>CY3008</td>
<td>Chirimoyos</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 14</td>
<td>EX4585</td>
<td>La Extensa</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 15</td>
<td>EX2668</td>
<td>La Extensa</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 16</td>
<td>EX2692</td>
<td>La Extensa</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 4</td>
<td>FN1583</td>
<td>San Francisco</td>
<td>Manabi</td>
<td>Sylvatic</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 5</td>
<td>TFN1678</td>
<td>San Francisco</td>
<td>Manabi</td>
<td>Sylvatic</td>
</tr>
<tr>
<td><em>P. chinai</em> 17</td>
<td>GA4354</td>
<td>Guara</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 18</td>
<td>GA4355</td>
<td>Guara</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 19</td>
<td>HW4820</td>
<td>Jaguay</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 20</td>
<td>HW4821</td>
<td>Jaguay</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 20</td>
<td>LE3346</td>
<td>La Encantada</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 21</td>
<td>LE2951</td>
<td>La Encantada</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 22</td>
<td>LE3349</td>
<td>La Encantada</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 21</td>
<td>LU4874</td>
<td>Lucarqui</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 23</td>
<td>MB2961</td>
<td>Maconta Abajo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 24</td>
<td>MB2958</td>
<td>Maconta Abajo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 22</td>
<td>ND2701</td>
<td>Naranjo Dulce</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 23</td>
<td>ND1730</td>
<td>Naranjo Dulce</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 6</td>
<td>PT2694</td>
<td>Pitayo</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 7</td>
<td>RN5775</td>
<td>San Ramon</td>
<td>Manabi</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 8</td>
<td>SC2810</td>
<td>El Sauce</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 24</td>
<td>SF2686</td>
<td>San Francisco</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 25</td>
<td>SG1346</td>
<td>San Gabriel</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 26</td>
<td>SG1348</td>
<td>San Gabriel</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>Location</td>
<td>Province</td>
<td>Type</td>
</tr>
<tr>
<td>----------------</td>
<td>------------</td>
<td>------------------</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td><em>P. chinai</em> 25</td>
<td>ST3013</td>
<td>Santa Ester</td>
<td>Loja</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. chinai</em> 26</td>
<td>ST4584</td>
<td>Santa Ester</td>
<td>Loja</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 9</td>
<td>TET2404</td>
<td>Estereo Seco</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 10</td>
<td>TET2388</td>
<td>Estereo Seco</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 11</td>
<td>THA2398</td>
<td>Humedad</td>
<td>Manabi</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 27</td>
<td>TKE2387</td>
<td>Pacoche</td>
<td>Manabi</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. rufotuberculatus</em> 12</td>
<td>TPY2392</td>
<td>Punta Larga</td>
<td>Manabi</td>
<td>Sylvatic</td>
</tr>
<tr>
<td><em>P. howardi</em> 28</td>
<td>TSW2317</td>
<td>Santa Rosa De Las Palmas</td>
<td>Manabi</td>
<td>Domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 29</td>
<td>ZL2778</td>
<td>Zapallo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 30</td>
<td>ZL2788</td>
<td>Zapallo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 31</td>
<td>ZL2790</td>
<td>Zapallo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 32</td>
<td>ZL2781</td>
<td>Zapallo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
<tr>
<td><em>P. howardi</em> 33</td>
<td>ZL2782</td>
<td>Zapallo</td>
<td>Manabi</td>
<td>Peri-domiciliary</td>
</tr>
</tbody>
</table>
Appendix 2.2: Preparation of plasmid extraction solutions.

**SOLUTION 1** (Kept at 4 °C)

2.5 mL 1M Tris pH 8.0
5 mL 0.2M EDTA
0.9g Dextrose (glucose)
Distilled water to 100 mL of Water
*Optional - Add RNASE to eliminate RNA.*

**SOLUTION 2**

Distilled water 8 mL
1 mL 10% SDS
1 mL 2N NaOH

**SOLUTION 3** (Kept at 4 °C)

80g Na Acetate-3H2O
50 mL Acetic Acid pH 4.5
Distilled water to 200 mL with water
Appendix 2.3: GARLI script used to conduct Cyt b phylogenetic inference using Maximum Likelihood method.

[generic]
datafname = CytbJan7Phyl72.nxs
constraintfile = none
streefname = stepwise
attachmentspertaxon = 50
ofprefix = CytbJanPhyl72run1.nuc.GTRIG
randseed = -1
availablememory = 512
logevery = 10
saveevery = 100
refinestart = 1
outputeachbettertopology = 0
outputcurrentbesttopology = 0
enforcetermconditions = 1
genthreshfortopoterm = 20000
scorethreshforterm = 0.05
significanttopochange = 0.01
outputphyliptree = 0
outputmostlyuselessfiles = 0
writecheckpoints = 0
restart = 0
outgroup = 1
searchreps = 2

datatype = nucleotide
ratematrix = 2rate
statefrequencies = estimate
ratehetmodel = gamma
numratecats = 4
invariantsites = none

[master]
nindivs = 4
holdover = 1
selectionintensity = 0.5
holdoverpenalty = 0
stopgen = 10000000
stoptime = 10000000
startoptprec = 0.5
minoptprec = 0.01
numberofprecreductions = 10
treerejectionthreshold = 50.0
topoweight = 1.0
modweight = 0.05
brlenweight = 0.2
randnniweight = 0.1
randsprweight = 0.3
limsprweight = 0.6
intervallength = 100
intervalstostore = 5

limsprrange = 6
meanbrlenmuts = 5
gammashapebrlen = 1000
gammashapemodel = 1000
uniqueswapbias = 0.1
distanceswapbias = 1.0

bootstrapreps = 1000
resampleproportion = 1.0
inferinternalstateprobs = 0
Appendix 2.4: GARLI script used to conduct ITS-2 phylogenetic inference using Maximum Likelihood method.

[general]
datafname = ITS-2Jan7Phyl72.nxs
constraintfile = none
streefname = stepwise
attachmentspertaxon = 50
ofprefix = ITS-2Jan7Phyl72run1.nuc.GTR1G
randseed = -1
availablememory = 512
logevery = 10
saveevery = 100
refinestart = 1
outputeachbettertopology = 0
outputcurrentbesttopology = 0
enforcetermconditions = 1
genthreshfortopoterm = 20000
scorethreshforterm = 0.05
significanttopochange = 0.01
outputphyliptree = 0
outputmostlyuselessfiles = 0
writecheckpoints = 0
restart = 0
outgroup = 1
searchreps = 2

datatype = nucleotide
ratematrix = 2rate
statefrequencies = estimate
ratehetmodel = none
numratecats = 1
invariantsites = none

[master]
nindivs = 4
holdover = 1
selectionintensity = 0.5
holdoverpenalty = 0
stopgen = 5000000
stoptime = 5000000

startoptprec = 0.5
minoptprec = 0.01
numberofprecreductions = 10
treerejectionthreshold = 50.0
topoweight = 1.0
modweight = 0.05
brlenweight = 0.2
randnniweight = 0.1
randsprweight = 0.3
limsprweight = 0.6
intervallength = 100
intervalstostore = 5

limsprrange = 6
meanbrlenmuts = 5
gammashapebrlen = 1000
gammashapemodel = 1000
uniqueswapbias = 0.1
distanceswapbias = 1.0

bootstrapreps = 1000
resampleproportion = 1.0
inferinternalstateprobs = 0
Appendix 2.5: MrBayes script used to conduct Cyt b phylogenetic inference using Bayesian method.

Model settings for partition 1:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Options</th>
<th>Current Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucmodel</td>
<td>4hy4/Doublet/Codon</td>
<td>4hy4</td>
</tr>
<tr>
<td>Nst</td>
<td>1/2/6</td>
<td>2</td>
</tr>
<tr>
<td>Code</td>
<td>Universal/Vertnt/Mycoplasma/</td>
<td>Metnt</td>
</tr>
<tr>
<td>Fstoe/Vertnt/Metnt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ploidy</td>
<td>Haploid/Diploid</td>
<td>Diploid</td>
</tr>
<tr>
<td>Rates</td>
<td>Equal/Gamma/Propinv/Inugamma/Adgamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>Nymacrat</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Nnetcat</td>
<td>&lt;number&gt;</td>
<td>5</td>
</tr>
<tr>
<td>Oneftp</td>
<td>Equal/Ny98/M3</td>
<td>Equal</td>
</tr>
<tr>
<td>Covarion</td>
<td>No/Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cading</td>
<td>All/Variable/Noabsencesites/Noabsencesites</td>
<td>All</td>
</tr>
<tr>
<td>Parmodel</td>
<td>No/Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

```
Mengv
This command sets the parameters of the Markov chain Monte Carlo (MCMC) analysis without actually starting the chain. This command is identical in all respects to Menu, except that the analysis will not start after this command is issued. For more details on the options, check the help menu for Mengv.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Options</th>
<th>Current Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed</td>
<td>&lt;number&gt;</td>
<td>1295818460</td>
</tr>
<tr>
<td>Sweepend</td>
<td>&lt;number&gt;</td>
<td>1295818460</td>
</tr>
<tr>
<td>Mgen</td>
<td>&lt;number&gt;</td>
<td>200000000</td>
</tr>
<tr>
<td>Mruns</td>
<td>&lt;number&gt;</td>
<td>2</td>
</tr>
<tr>
<td>Mchains</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Temp</td>
<td>&lt;number&gt;</td>
<td>0.200000</td>
</tr>
<tr>
<td>Reweight</td>
<td>&lt;number&gt;</td>
<td>0.00 u 0.00 ^</td>
</tr>
<tr>
<td>Swepfreq</td>
<td>&lt;number&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Neastay</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Samplerefq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Printfreq</td>
<td>&lt;number&gt;</td>
<td>1000</td>
</tr>
<tr>
<td>Pintall</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Printmax</td>
<td>&lt;number&gt;</td>
<td>9</td>
</tr>
<tr>
<td>Membdiag</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Diggrefq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Minsumfreq</td>
<td>&lt;number&gt;</td>
<td>0.10</td>
</tr>
<tr>
<td>Allchains</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Allcomps</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Rburnin</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Burnin</td>
<td>&lt;number&gt;</td>
<td>0.25</td>
</tr>
<tr>
<td>Stopple</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Stopval</td>
<td>&lt;number&gt;</td>
<td>0.01</td>
</tr>
<tr>
<td>Filename</td>
<td>&lt;name&gt;</td>
<td>CytbJanPhyML72.nxs.&lt;p/t&gt;</td>
</tr>
<tr>
<td>Startingtree</td>
<td>Random/User</td>
<td>Random</td>
</tr>
<tr>
<td>Npart</td>
<td>&lt;number&gt;</td>
<td>0</td>
</tr>
<tr>
<td>Sheehrsense</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ordertaxa</td>
<td>Yes/No</td>
<td>No</td>
</tr>
</tbody>
</table>
```
Appendix 2.6: MrBayes script used to conduct ITS-2 phylogenetic inference using Bayesian method.

### Model settings for partition 1:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Options</th>
<th>Current Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucmodel</td>
<td>4by4/Doublet/Codon</td>
<td>4by4</td>
</tr>
<tr>
<td>Nst</td>
<td>1/3/6</td>
<td>6</td>
</tr>
<tr>
<td>Code</td>
<td>Universal/Vertm/Mycoplasma/Yeast/Ciliates/Metnt</td>
<td>Universal</td>
</tr>
<tr>
<td>Ploidy</td>
<td>Haploid/Diploid</td>
<td>Diploid</td>
</tr>
<tr>
<td>Rates</td>
<td>Equal/Gamma/Propinv/Invgamma/Adgamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>Ngammacat</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Nhetacat</td>
<td>&lt;number&gt;</td>
<td>5</td>
</tr>
<tr>
<td>OmegaPsi</td>
<td>Equal/Ny98/M3</td>
<td>Equal</td>
</tr>
<tr>
<td>Covarion</td>
<td>No/Yes</td>
<td>No</td>
</tr>
<tr>
<td>Coding</td>
<td>All/Variable/Noabsencesites/Nonpresencesites</td>
<td>All</td>
</tr>
<tr>
<td>Parsmodel</td>
<td>No/Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

### Mecmep

This command sets the parameters of the Markov chain Monte Carlo (MCMC) analysis without actually starting the chain. This command is identical in all respects to Mecm, except that the analysis will not start after this command is issued. For more details on the options, check the help menu for Mecm.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Options</th>
<th>Current Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed</td>
<td>&lt;number&gt;</td>
<td>1295818460</td>
</tr>
<tr>
<td>Swapseed</td>
<td>&lt;number&gt;</td>
<td>1295818460</td>
</tr>
<tr>
<td>Ngen</td>
<td>&lt;number&gt;</td>
<td>2000000000</td>
</tr>
<tr>
<td>Nruns</td>
<td>&lt;number&gt;</td>
<td>2</td>
</tr>
<tr>
<td>Nchains</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Temp</td>
<td>&lt;number&gt;</td>
<td>0.2000000</td>
</tr>
<tr>
<td>Reweight</td>
<td>&lt;number&gt;</td>
<td>0.00 u 0.00</td>
</tr>
<tr>
<td>Swapfreq</td>
<td>&lt;number&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Switch</td>
<td>&lt;number&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Samplefreq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Printfreq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Printall</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Printmax</td>
<td>&lt;number&gt;</td>
<td>8</td>
</tr>
<tr>
<td>Mcmediagn</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Diagnfreq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Minpartfreq</td>
<td>&lt;number&gt;</td>
<td>0.10</td>
</tr>
<tr>
<td>Allchains</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Allcomps</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Reburnin</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Burnin</td>
<td>&lt;number&gt;</td>
<td>0</td>
</tr>
<tr>
<td>Burninfrac</td>
<td>&lt;number&gt;</td>
<td>0.25</td>
</tr>
<tr>
<td>Stopfile</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Stopval</td>
<td>&lt;number&gt;</td>
<td>0.01</td>
</tr>
<tr>
<td>Filename</td>
<td>Name.</td>
<td>ITS-2-Jan?Phy172.nxs.p/t</td>
</tr>
<tr>
<td>Startingtree</td>
<td>Random/User</td>
<td>Random</td>
</tr>
<tr>
<td>Users</td>
<td>&lt;number&gt;</td>
<td>0</td>
</tr>
<tr>
<td>Savehlength</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ordertaxa</td>
<td>Yes/No</td>
<td>No</td>
</tr>
</tbody>
</table>
Appendix 2.7: *GARLI* script used to conduct phylogenetic inference of South American ITS-2 samples using Maximum Likelihood method.

```plaintext
[general]
datafname = ITS-2SAMPLES.nxs
constraintfile = none
streefname = stepwise
attachmentspertaxon = 50
ofprefix = ITS-2SAMPLESFeb18run1.nuc.GTRIG
randseed = -1
availablememory = 512
logevery = 10
saveevery = 100
refinestart = 1
outputeachbettertopology = 0
outputcurrentbesttopology = 0
enforcetermconditions = 1
genthalphfortopoterm = 20000
scorethreshforterm = 0.05
significanttopochange = 0.01
outputphyliptree = 0
outputmostlyuselessfiles = 0
writecheckpoints = 0
restart = 0
outgroup = 1
searchreps = 2

datatype = nucleotide
ratematrix = 6rate
statefrequencies = estimate
ratehetmodel = gamma
numratecats = 4
invariantsites = none

[master]
nindivs = 4
holdover = 1
selectionintensity = 0.5
holdoverpenalty = 0
stopgen = 5000000
stoptime = 5000000

startoptprec = 0.5
minoptprec = 0.01
numberofprecreductions = 10
treerejectionthreshold = 50.0
topoweight = 1.0
```
modweight = 0.05
brlenweight = 0.2
randnniweight = 0.1
randsprweight = 0.3
limsprweight = 0.6
intervallength = 100
intervalstostore = 5

limsprrange = 6
meanbrlenmuts = 5
gammashapebrlen = 1000
gammashapemodel = 1000
uniqueswapbias = 0.1
distanceswapbias = 1.0

bootstrapreps = 1000
resampleproportion = 1.0
inferinternalstateprobs = 0
Appendix 2.8: MrBayes script used to conduct phylogenetic inference of South American ITS-2 samples using Bayesian method.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Options</th>
<th>Current Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mucmodel</td>
<td>4by4/Doublet/Codon</td>
<td>4by4</td>
</tr>
<tr>
<td>Rate</td>
<td>1/2/5</td>
<td>6</td>
</tr>
<tr>
<td>Code</td>
<td>Universal/Vertmic/Mycoplasma/</td>
<td>Universal</td>
</tr>
<tr>
<td>Yeast/Ciliates/Metmt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ploidy</td>
<td>Diploid/Diploid</td>
<td>Diploid</td>
</tr>
<tr>
<td>Rates</td>
<td>Equal/Gamma/Propino/Inugamma/Adgamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>NgammaCat</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Whetcat</td>
<td>&lt;number&gt;</td>
<td>5</td>
</tr>
<tr>
<td>Onew 90%</td>
<td>Equal/My98/M3</td>
<td>Equal</td>
</tr>
<tr>
<td>Covarion</td>
<td>No/Yes</td>
<td>No</td>
</tr>
<tr>
<td>Coding</td>
<td>All/Variable/Noabsencesites/</td>
<td>All</td>
</tr>
<tr>
<td>Parsmodel</td>
<td>Nopresencesites</td>
<td></td>
</tr>
</tbody>
</table>

**Mcmcp**

This command sets the parameters of the Markov chain Monte Carlo (MCMC) analysis without actually starting the chain. This command is identical in all respects to Mmc, except that the analysis will not start after this command is issued. For more details on the options, check the help menu for Mmc.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Options</th>
<th>Current Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed</td>
<td>&lt;number&gt;</td>
<td>1390119524</td>
</tr>
<tr>
<td>Swapseed</td>
<td>&lt;number&gt;</td>
<td>1390119524</td>
</tr>
<tr>
<td>Ngen</td>
<td>&lt;number&gt;</td>
<td>100000000</td>
</tr>
<tr>
<td>Nruns</td>
<td>&lt;number&gt;</td>
<td>2</td>
</tr>
<tr>
<td>Nchains</td>
<td>&lt;number&gt;</td>
<td>4</td>
</tr>
<tr>
<td>Temp</td>
<td>&lt;number&gt;</td>
<td>0.2000000</td>
</tr>
<tr>
<td>Nweight</td>
<td>&lt;number&gt;,&lt;number&gt;</td>
<td>0.00 v 0.00</td>
</tr>
<tr>
<td>Swaptopf</td>
<td>&lt;number&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Nswaps</td>
<td>&lt;number&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Samplefreq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Printfreq</td>
<td>&lt;number&gt;</td>
<td>1000</td>
</tr>
<tr>
<td>Printall</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Printmax</td>
<td>&lt;number&gt;</td>
<td>9</td>
</tr>
<tr>
<td>Mcmctopf</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Diagnfreq</td>
<td>&lt;number&gt;</td>
<td>10000</td>
</tr>
<tr>
<td>Minpartfreq</td>
<td>&lt;number&gt;</td>
<td>0.10</td>
</tr>
<tr>
<td>Allchains</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Allcomps</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>ReBurnin</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Burnin</td>
<td>&lt;number&gt;</td>
<td>0</td>
</tr>
<tr>
<td>Burninfrac</td>
<td>&lt;number&gt;</td>
<td>0.25</td>
</tr>
<tr>
<td>Stoprule</td>
<td>Yes/No</td>
<td>No</td>
</tr>
<tr>
<td>Stopval</td>
<td>&lt;number&gt;</td>
<td>0.91</td>
</tr>
<tr>
<td>Filename</td>
<td>(name)</td>
<td>ITS-2SAMPLESFeb18.nex.&lt;p/t&gt;</td>
</tr>
<tr>
<td>Startingtree</td>
<td>Random/User</td>
<td>Random</td>
</tr>
<tr>
<td>Nperts</td>
<td>&lt;number&gt;</td>
<td>0</td>
</tr>
<tr>
<td>Savebrlns</td>
<td>Yes/No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ordertaxa</td>
<td>Yes/No</td>
<td>No</td>
</tr>
</tbody>
</table>