POPULATION GENETICS AND NATURAL HISTORY OF *TRIATOMA SANGUISUGA* IN SOUTHEASTERN LOUISIANA

A DISSERTATION

SUBMITTED ON THE TENTH DAY OF MARCH 2017

TO THE DEPARTMENT OF TROPICAL MEDICINE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF PUBLIC HEALTH AND TROPICAL MEDICINE

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

BY

Samuel B. Jameson

APPROVED:_____

Dawn Wesson, PhD **Date**

_____

Zeynep Özsoy, PhD **Date**

_____

Juan Pizarro, PhD **Date**

_____

Sunshine Van Bael, PhD **Date**

Acknowledgements

I would like to thank Dr. Dawn Wesson for her perennial support of my studies. She has provided a steady guiding hand throughout my academic and professional development. In all likelihood, I would not have pursued research without her insistence. Likewise, I would not have grown academically or professionally without her corrections where they were needed. I will forever be grateful for her mentorship. Additionally, I would like to thank the other members of my committee: Dr. Zeynep Özsoy, Dr. Juan Pizarro, and Dr. Sunshine Van Bael. I would like to thank Dr. Özsoy for being kind enough to lend me laboratory time and her expertise while I was in Colorado. She provided invaluable help with DNA sequencing, phylogenetics, and phylogenetic markers. In a region where academic research is scarce, Dr. Özsoy was extremely helpful in maintaining the momentum of and my active interest in this research. I would like to thank Dr. Pizarro for all of the help that he has provided me in my capacity as a student and as a staff member. He was always available to lend his expertise in molecular methods, and his even keeled disposition has been an inspiration and a model for me professionally. Lastly, I would like to thank Dr. Van Bael for her critical eye. Her questions and corrections through this process have been vital to creating a cogent and reasoned dissertation. Her reminders that science is not good science without a sound and defined structure have been indispensable while constructing this dissertation.

I would like to thank the many students and fellows that have helped me with laboratory methods over the years: Dr. Mark Rider, Karen Boegler, Dr. Robin Moudy, Dr. Justin Davis, and Dr. Michael Deshotel. I would like to especially thank Dr. Claudia Herrera for her years of guidance and friendship. Her passion for *Trypanosoma cruzi*

research and willingness to share methods and ideas has been a true inspiration and vital to my growth in this field of study. I would also like to thank Dr. Mike Garrett at the University of Mississippi Medical Center who donated untold hours of machine time to complete this research.

I would like to thank the many student who have participated in the field collection and nucleic acid extractions utilized in the course of this research: Velma Lopez, Jessica Saeger, Karen Boegler, Michelle Blyth, Dr. Poppy Markwell, and many summer student workers and volunteers. I would like to especially thank Sarah Michaels and Erin Cloherty for their help with several years of insect collections and friendship along the way. My training and graduate studies have followed Sarah's closely; as such, she has been a consistent sounding board for me throughout my studies. Erin accompanied me in the majority of the field collections used in these studies. Her friendship and infectious up-beat personality has made many hours of summer field collections a joy.

I would like to thank the individuals and groups that have supported me over the years: the laboratory of Dr. Wesson through grants from the Bill and Melinda Gates Foundation; Dr. Pierre Buekens – Dean of Tulane's School of Public Health and Tropical Medicine; Mr. H. Merritt Lane, III; Dr. Preston Marx and the Department of Tropical Medicine; and Dr. Charlotte Hobbs at the University of Mississippi Medical Center.

I would like to thank my family and friends for their love, support, and patience throughout this process. My parents, Drs. Mike and Phyllis Nowicki, have championed my studies, encouraged me, and supported me constantly. Erin, who opened her home to me, has made the completion of this work logistically possible. Finally, I would like to

thank my husband, Andrew Jameson. He has been my rock and my biggest fan for more than seven years. This dissertation is only possible because of his love and support.

Table of Contents

Abstract

The United States is home to eleven species of triatomine insects that can carry and transmit the parasite that causes Chagas disease, *Trypanosoma cruzi*. Much of the natural history of these vectors remains unknown. This is especially true of the species native to southeastern Louisiana, *Triatoma sanguisuga*. To this end experiments were undertaken to understand the movements, blood meal sources, and associated vertebrates of *T. sanguisuga*. Sites of genetic variability were sequenced and compared to determine how related *T. sanguisuga* spatially aggregate. It is demonstrated that the cytochrome b gene can be used to identify sub-populations of these insects and monitor their spatial aggregation over a multi-year period. Genetic analysis suggests the presence of wide-spread sub-populations. The sources of vertebrate blood meals were determined from collected specimens to determine the most frequent bloodmeal sources. The most frequent bloodmeal sources were found to be from humans, cows, pigs, dogs, and raccoons - though some of these sequences are likely the result of laboratory contamination. Finally, habitat modeling of *T. sanguisuga* provides a framework for assessing the shared habitats of *T. sanguisuga* and its vertebrate food sources. By using publicly available reports of *T. sanguisuga*, a model representing suitable habitat for the vector was created. By comparing this distribution to similar distribution models of associated animals, it was determined that the most statistically significant overlap of suitable habitats occurs between *T. sanguisuga* and two small rodent species, *Oryzomys palustris* and *Neotoma floridana*. The present research adds significantly to the limited knowledge of the natural history of this vector species. The work provides a foundation

on which to build evidence-based risk assessments of autochthonous human Chagas disease in southeastern Louisiana and beyond.

List of Tables

List of Figures

Chapter I

Background and Significance

The protozoan parasite *Trypanosoma cruzi* is the etiologic agent of human American trypanosomiasis, also known as Chagas disease. World Health Organization (WHO) estimates that *T. cruzi* infects 5-6 million people in Central and South America. Additionally, the WHO considers another 25 million people to be at risk of infection [1]. Although native to the Americas, the prevalence of Chagas disease is increasing around the world. This spread is the result of immigration and ever-easier international travel. It has been estimated that there are greater than 300,000 non-native *T. cruzi* infected individuals in the United States and more than 5,500 such individuals in Canada. Other regions that have had noticeable increases in non-native Chagas disease cases include Europe, Japan, and Australia [2]. These figures ignore the fact that a substantial portion of the United States maintains a very low level of autochthonous Chagas disease. Unfortunately, there is currently no research quantifying the risk of autochthonous Chagas disease in the southern United States where there are both competent arthropod vectors and an abundance of *T. cruzi* [3–6]. Current detection and reporting of autochthonous Chagas disease in the United States is based on routine, but not universal, screening of donated blood [7].

The classic arthropod vectors of *T. cruzi* are all Hemipterans in the family Reduviidae and the subfamily Triatominae. These triatomine bugs have various colloquial names across their natural range including kissing bugs, vinchuca, and barbeiros. There are at least 130 recognized species of triatomine bugs and many of these are competent vectors of *T. cruzi* and feed readily on vertebrates [8]. However, fewer

than twenty of these species are thought to be responsible for nearly all Chagas disease [9]. In the United States, there are 11 species of triatomine bugs – all of which are competent vectors of *T. cruzi*. The two species that have been reported in Louisiana are *Triatoma lecticularia* and *Triatoma sanguisuga* [10]. The only species reported in southeastern Louisiana is *T. sanguisuga* [11]. *T. sanguisuga* (LeConte) is also known by its common name, the eastern bloodsucking conenose. It has been reported throughout the southeastern United States

It is considered a sylvatic or peridomestic species [8]. To date, there have been no reports of *T. sanguisuga* successfully colonizing a human residence built with modern construction techniques. However, adult *T. sanguisuga* may invade homes during the summer months. This is facilitated by the ability of adult triatomine bugs to fly. The nymphal instars of all triatomine bugs are wingless and, thus, have a limited dispersal range [5]. Anecdotally, bright lights in or around a home have been found to be associated with the presence of *T. sanguisuga* in the home. This attraction has been studied in a closely related species, *Triatoma dimidiata*, in the Yucatan peninsula of Mexico [12]. In a previous study, the additional factors of cats, chickens, and the lack of air conditioning were found to be positively correlated with *T. sanguisuga* presence at a home [13]. It is important to note that *T. sanguisuga* has not been reported in urban or suburban areas in Louisiana. Based on locations at which *T. sanguisuga* have been found in southeastern Louisiana, it is thought that a heavily wooded area is needed to support a population of these bugs. Human interaction with these bugs seems to happen incidentally when adults are attracted to homes. While this attraction increases the likelihood of human-bug interactions and, consequently, *T. cruzi* infection in humans,

autochthonous Chagas disease cases remain extremely rare in the United States [7]. This remains and interesting and active area of research because more than half of the field collected *T. sanguisuga* in the region have been found to carry the parasite [3, 14]. By analyzing survey data, researchers have described several possible risk factors for autochthonous *T. cruzi* infection in the United States. Camping without a tent during summer months and a history of exposure to small game animals were found to be the most significant risk factors [15].

In the typical scenario of vector-borne transmission of *T. cruzi* to humans, triatomine bugs emerge from hiding at night to feed on sleeping humans. As the bugs' midguts fill with blood, the abdomen expands and the contents of the hindgut are expelled onto the skin near the bite. The irritation caused by the bite or sensation of the bug walking on the person's skin causes reflexive scratching of the site. It is this scratching that pushes the parasite-laden excreta into the bite wound or onto a mucosal membrane and allows infection [16]. Anecdotal observations have suggested that an extended feeding-to-defecation time compared to more efficient vectors (e.g. *Triatoma infestans*) is largely responsible for *T. sanguisuga*'s low transmission efficiency. However, defecation during feeding has been observed on several instances in the laboratory maintenance of *T. sanguisuga* indicating more research is needed on this aspect of *T. sanguisuga* biology.

T. cruzi is a kinetoplastid parasite of humans and many other warm-blooded vertebrates. The metacyclic trypomastigote form of *T. cruzi* is the highly infectious form passed from the rectum of the triatomine bug to the vertebrate host through the bite wound left by the bug. Metacyclic trypomastigotes are not able to penetrate an intact

epidermis, but they can readily cross the intact mucosal membranes of the mouth and conjunctiva – providing a secondary means of infection. After entering the blood stream of the vertebrate host, the trypomastigotes invade healthy cells and differentiate into the reproductive amastigote form. The amastigotes reproduce by several rounds of binary fission within the cytoplasm before transforming back into trypomastigotes and being released by cell lysis to infect more cells. Peripherally circulating trypomastigotes are ingested by feeding triatomine bugs where they begin the second phase of their life cycle. In the midgut of the bug, *T. cruzi* differentiate into a second replicative form – the epimastigote. The epimastigotes undergo many rounds of binary fission before they pass into the hindgut and rectum. In the rectum, these parasites differentiate into the infectious metacyclic trypomastigote form to continue the cycle (Figure 2).

Figure 1 - Suspected range of *T. sanguisuga* based on occurrence reports aggregated by the Centers for Disease Control and Prevention. It remains unclear why there are no reports of *T. sanguisuga* in West Virginia [97].

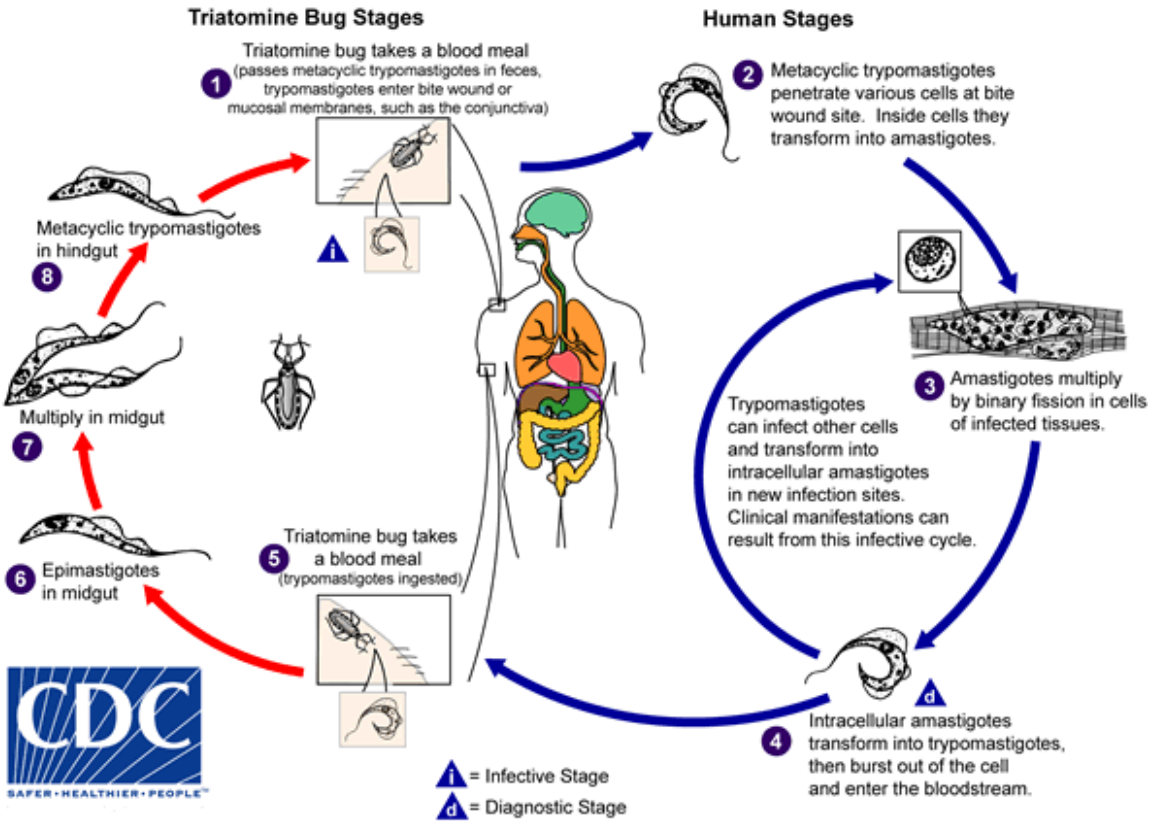


Figure 2 - Lifecycle of *T. cruzi*. Although only the classical system of vector-born transmission is outlined in the graphic, other important routes of infection include entry through the conjunctiva of the eye and ingestion [99].

T. cruzi is currently separated into six discrete typing units (DTUs) according to international consensus [17]. It has been suggested that each DTU (codified as: TcI – TcVI) exists within finite geographic areas and is maintained in specific, but often overlapping cycles (Figure 3) [18]. This idea has been challenged on some level by the recent detection of TcII in a sylvatic rodent in Orleans Parish [8]. Previously, TcII was thought to be geographically confined to areas south of the Amazon rainforest [18]. These findings are supported by recent detection of many DTUs once thought exotic in vector populations in Veracruz, Mexico [19].

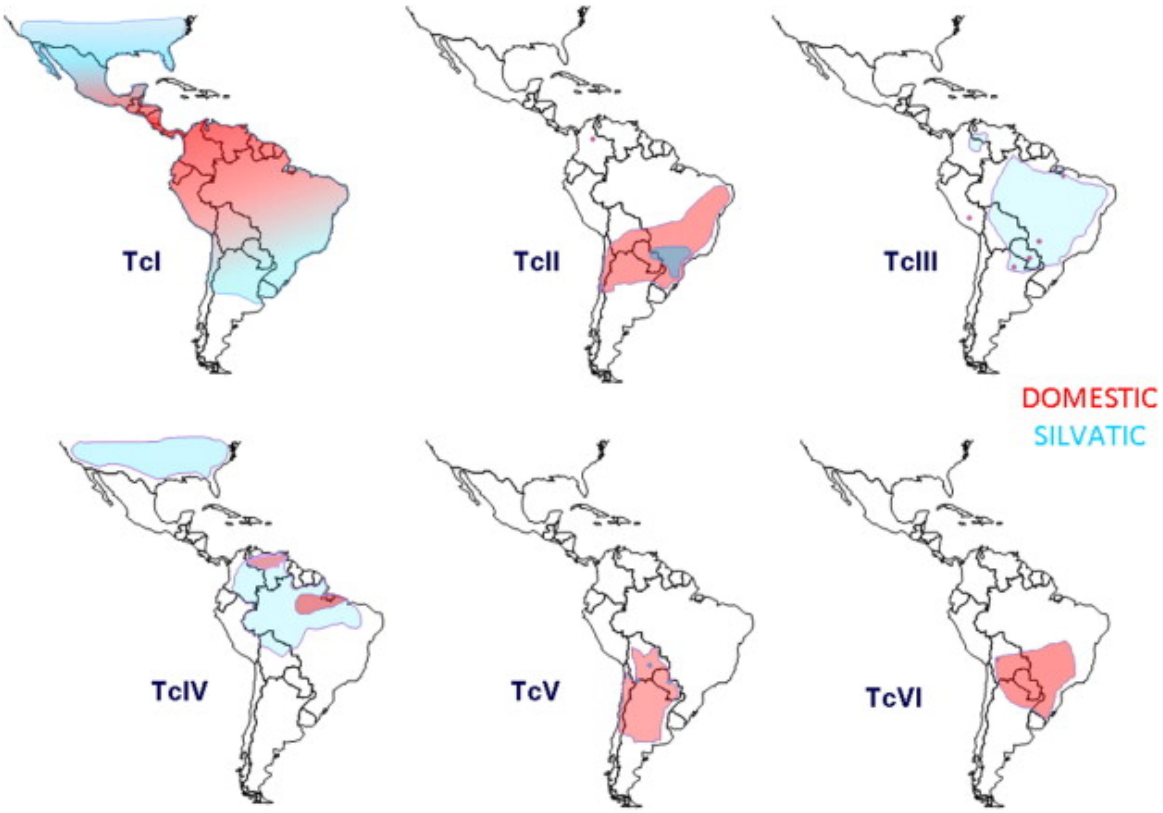


Figure 3 - Distribution of discrete typing units of *T. cruzi*. Although these distributions remain widely accepted, recent research is challenging the assumption of such clear delineations of parasite populations. [18].

Chapter II

Literature Review, Rationale, and Approach

There remains a great deal that is unknown about the population structure and natural history of *T. sanguisuga*. The foundational research on the ecology of *T. sanguisuga* and other triatomines of the United States took place in the 1940's and 1950's [20–23]. Unfortunately, triatomine taxonomy had not been fully agreed upon or standardized at that time. The literature continues to suffer from these earlier misclassifications of triatomine species. One notable example is the Usinger's 1944 paper that clearly differentiated *T. lecticularia* from *T. sanguisuga* for the first time while simultaneously claiming the range of *T. sanguisuga* extends west to southern California [24]. The accounts of *T. sanguisuga* in California are now disregarded entirely as the farthest extent of its western range has been found to be Texas [25]. Such long-standing confusion in the literature necessitates comprehensive examination and correction. Although an authoritative review of research involving *T. sanguisuga* falls outside of the scope of the present work, several disputed aspects of the natural history of *T. sanguisuga* have been identified for further study including the frequency with which it moves through space, its feeding habits, and the animals with which it is associated.

Large- and small-scale population genetics of *T. sanguisuga*

Only one previous study has been undertaken to genetically demonstrate the large- and small-scale distributions of *T. sanguisuga* [26]. This study used the cytochrome b and 16S rDNA genes to make phylogenetic inferences from a small sample of *T. sanguisuga* specimens. Both of these mitochondrial genes are commonly used for phylogenetic studies. The cytochrome b gene encodes an important protein in the electron transport chain, while the 16S gene encodes part of the small subunit of ribosomes. Because of their importance in sustaining life, mutations in these genes accumulate predictable and well defined regions that do not negatively impact the function of the gene. By amplifying these regions of mitochondrial DNA, inferences can often be made about the population structure, evolution, and divergence of closely related species. The utility of these markers has been demonstrated in other triatomine vectors such as *Rhodnius prolixus*, *Rhodnius robustus,* and *Rhodnius nasutus* [27]. Because there has only been one study attempting to characterize the population genetics of *T. sanguisuga* at small and large geographic scales, it is important to ensure that the results are reproducible and hold across a large sample size. Indeed, the results reported from this study [26] are unexpected given what little is known of the natural history of this vector. Briefly, the researchers examined genetic sequences from fifty-four *T. sanguisuga* specimens collected over two years from a single location. From these specimens, they detected thirty-eight haplotypes that segregated into two groups with significant bootstrap support in both neighbor-joining and maximum parsimony based analysis. Various statistical methods including analysis of molecular variance (AMOVA), Fu's *F*, and Chakraborty's amalgamation testes were used to determine the stability of the genetic

loci within the population. The separate phylogenetic groups were identified both in the primary collection site and in two *T. sanguisuga* samples from greater than sixty miles away. With these data, the researchers suggest that what appears to be a single species of *T. sanguisuga* may, in fact, be two or more sibling species inhabiting the same geographic area. It is believed that neither of the recognized subspecies of *T. sanguisuga* (*T. s. texana* and *T. s. ambigua*) has ever been reported in Louisiana.

Major limiting factors to the previous research are the small sample size, the short duration of the bug collections, and a lack of high-resolution spatial data of the collection location of each bug. A robust sample size is helpful for any study of population genetics, but it becomes imperative in situations where there is no reference genome assembly and genetic diversity is larger than expected. By increasing the sample size, one would be able to obtain a more accurate assessment of different breeding populations or subspecies in a given area. Furthermore, the collections in this study were performed over two consecutive summers. While this would be a more than sufficient timeframe for many arthropods, the reproductive rate for *T. sanguisuga* is quite slow. *T. sanguisuga* take roughly 12 months to grow from an egg to a sexually mature adult [23]. This means that the collections in the study were, effectively, a cross-section of the existing genetic diversity. By extending the time over which collections are made, it is more likely that one would see statistically significant changes in genotype groups over the years. Finally, because it is thought that the flight range of *T. sanguisuga* is limited [28], high-resolution data about their collection site could potentially be useful in separating haplotypes spatially. If genetic groups were clearly separated by a physical distance, it would lend support to the hypothesis of separate breeding groups or even subspecies.

The present research aims to verify the utility of the cytochrome b gene of *T. sanguisuga* in determining sub-groups of *T. sanguisuga* across small and large distances. A larger sample size (both temporally and spatially) is used to increase the validity and resolution of the conclusions. Additionally, fine-resolution spatial data is used to segregate genetic groups of *T. sanguisuga* collected from the same location as the fifty-four samples previously mentioned [26]. By utilizing the same genetic locus, one can build on the previous study to further expand upon these poorly understood genetic relationships.

Animals associated with *T. sanguisuga*

Many triatomine vector species have a primary animal with which they are associated (Table 1). These animals all create nests, middens, or burrows. Stationary residences are an important characteristic of animals associated with triatomine vectors. There are four primary reasons that a stationary residence is required for triatomines: immature stages cannot fly, a blood meal is required for each molt, triatomines typically take seven to twelve months to reach the adult stage, and triatomines are adapted for feeding on sleeping animals. Taken together, the biological constraints necessitate that triatomines spend much of their life in close proximity to the sleeping location of a suitable blood meal source [1]. Analysis of the blood meal or residual DNA from a blood meal is a powerful tool to trace the movements and interactions of a triatomine [29]. Unfortunately, this kind of blood meal analysis can lead to confusing and biased results. There are two major sources of bias in the current literature regarding blood meal analysis from triatomines, in general, and *T. sanguisuga*, specifically.

The primary source of bias is the developmental stage of the sampled insect. Common methods of collection for *T. sanguisuga* include passive surveillance of homes and nighttime white sheet collections. Both methods exclusively target the flying adult stage. This is problematic when attempting to determine what, if any, animal is the primary associated mammal of *T. sanguisuga*. Blood meal analysis of triatomines in North America have relied heavily on these adult-biased collection methods [29–32]. Determining the blood meal source from adults is important for public health reasons [32]. However, knowledge of the natural history and primary associated mammal of *T. sanguisuga* would allow for targeted control efforts of the immature stages. A common

interpretation of blood meal analysis data from adult specimens is to label every identified blood meal source as an associated animal of *T. sanguisuga* [30]. Additionally, any animal testing positive for *T. cruzi* within the reported distribution of *T. sanguisuga* is referred to as an associated animal of *T. sanguisuga* [33]. These practices have resulted in an ever-expanding roster of animals that are reportedly associated with this species (Table 2 and Table 3). Previous blood meal analyses of *T. sanguisuga* in the literature have had sample sizes of less than fifty insects collected from only a few locations [29, 32]. To obtain a comprehensive picture of adult blood meal sources, one would need to analyze hundreds of insects over a large geographic distance.

The secondary source of bias in interpreting triatomine blood meal analysis results is the assumption that detecting vertebrate DNA in a given sample means that the insect fed on that vertebrate. Investigators in several countries have fallen prey to this assumption with regard to human DNA [32, 34, 35]. What is usually omitted from these reports is the possibility of contamination. Human DNA is a common contaminant of consumables and surfaces in research laboratories. Indeed, popular pipette tips like ART™ Barrier tips are labeled for “PCR” or “Molecular” use, but they are not certified free of human DNA. One such paper [35] targeted only sylvatic bugs that were assumed to have very little human contact. However, twenty-seven of the ninety-eight samples were found to contain human DNA. Most surprisingly, sixty-six percent of these human samples were in second to fifth instar nymphs, which they acknowledge have a limited dispersal range. The conclusion reached by these and other researchers is that triatomine species, including *T. sanguisuga,* that are thought to be sylvatic feed on humans with significant frequency – thus, posing an increased public health threat. Only recently has a

single publication mentioned and attempted to control for the possibility of human DNA contamination in their samples [29]. Although they were not able to show that the human DNA was contamination, they were able to demonstrate that the samples were collected from homes where they could have fed on humans.

The present research seeks to build upon the existing body of knowledge regarding the blood meal sources of *T. sanguisuga*. To control for possible biases in collection and interpretation of data, a large sample size including samples from across southeastern Louisiana is analyzed. Additionally, samples that are identified as human DNA are haplotyped and compared to DNA samples from the individuals that originally extracted the DNA from the insect. By haplotyping the human DNA and knowing the location at which the insect was collected, one can identify human DNA as known contamination, likely contamination, or likely blood meal source.

Table 1 - Important vectors of *T. cruzi* and their primary associated animals.

Triatomine species	**Primary associate animal(s)**	**Citation**
Triatoma infestans	Humans; Common Yellow-toothed Cavy	[36]
Triatoma brasiliensis	Small, rock-dwelling rodents	[37]
Panstrongylus megistus	Nesting birds, burrowing rodents	[38]
Rhodnius prolixus	Palm tree-dwelling birds, rodents, and marsupials	[39]
Triatoma dimidiata	Nesting birds, opossums, and rodents	[40]
Panstrongylus geniculatus	*Rattus rattus* (urban); armadillo and opossum (sylvatic)	[41, 42]
Triatoma rubrofasciata	Humans, *Rattus rattus*	[42]

Table 2 - Vertebrate blood meals identified from *T. sanguisuga*.

Common Name	**Scientific Name**	**Citation**
Eastern woodrat	*Neotoma floridana**	[32]
Raccoon	*Procyon lotor*	[32]
Domestic cat	*Felis catus*	[32]
Frog	Not speciated	[32]
Human	*Homo sapiens*	[5, 32]
Cow	*Bos taurus*	[32]
Domestic dog	*Canis familiaris*	[5, 32]
Squirrel	*Sciurus carolinensis* or *Sciurus niger**	[32]

*Probable scientific names based on location of insect collection and documented species in that region.

Table 3 - Animals associated with *T. sanguisuga* based on infection with *T. cruzi*.

Common Name	**Scientific Name**	**Citation**
Eastern woodrat	*Neotoma floridana*	[8, 23]
Cotton mouse	*Peromyscus gossypinus*	[8]
Horse	*Equus ferus*	[23]
Western cotton rat	*Sigmadon hispidus*	[23]
Chicken	*Gallus gallus domesticus*	[23]
Human	*Homo sapiens*	[23, 43]
Domestic dog	*Canis familiaris*	[44]
Virginia opossum	*Didelphis virginiana*	[6, 45]
Raccoon	*Procyon lotor*	[6, 45]
Tree toads	*Hyla* spp.	[43]
Woodrat	*Neotoma* spp.	[46]
Nine-banded armadillo	*Dasypus novemcinctus*	[47]
Domestic cat	*Felis catus*	[13]
Lion-tailed macaque	*Macaca silenus*	[33]
Ring-tailed lemur	*Lemur catta*	[33, 48]
Pig-tailed macaque	*Macca nemestrina*	[48]
Rhesus macaque	*Macca mulatta*	[49]
Baboon	*Papio* spp.	[49]
Pileated gibbon	*Hylobates pileatus*	[50]
Black and white ruffed lemur	*Varecia variegate variegata*	[48]
Black-eyed lemur	*Eulemur macaco flavifrons*	[48]
Striped skunk	*Mephitis mephitis*	[45]
Bobcat	*Lynx rufus*	[45]
Coyote	*Canis rufus*	[45]

Ecological niche modeling of *T. sanguisuga*

The final disputed aspect of the natural history of *T. sanguisuga* that the present research aims to clarify is the question of ecological niche and species distribution. As previously mentioned, researchers often discuss the ecologies of triatomine vectors with relation to an associated animal. The consensus in the literature is that triatomine vectors behave in much the same way as their relative, the bed bug (*Cimex lectularius*) [51]. That is to say, they remain tightly associated with a stationary animal for much of their lives and have a limited dispersal range [1, 23]. Unlike bed bugs, adult triatomines have wings and can fly short distances [28]. However, because triatomines take many months to mature to adulthood and require blood meals for each of their five molts, the role of a given triatomine's associated animal is significant. While most of the triatomine vectors in the United States have a known associated animal or a well-defined habitat, *T. sanguisuga* is reported to have many associated animals and no defined ecological niche [25]. The currently proposed distributions of *T. sanguisuga* are either broadly based on state-level reported collections of live *T. sanguisuga* (Figure 4) or are ecological niche models based on limited data sets (Figure 5). While the former method is informational at the national level, it is overly broad and does not differentiate amongst varying habitats and environments within states. The latter is based on a limited number of samples (n=18). This small sample is not representative of the true distribution of *T. sanguisuga* as live specimens are routinely collected in southeastern Louisiana in areas that this model deems inhospitable for the species [3, 26].

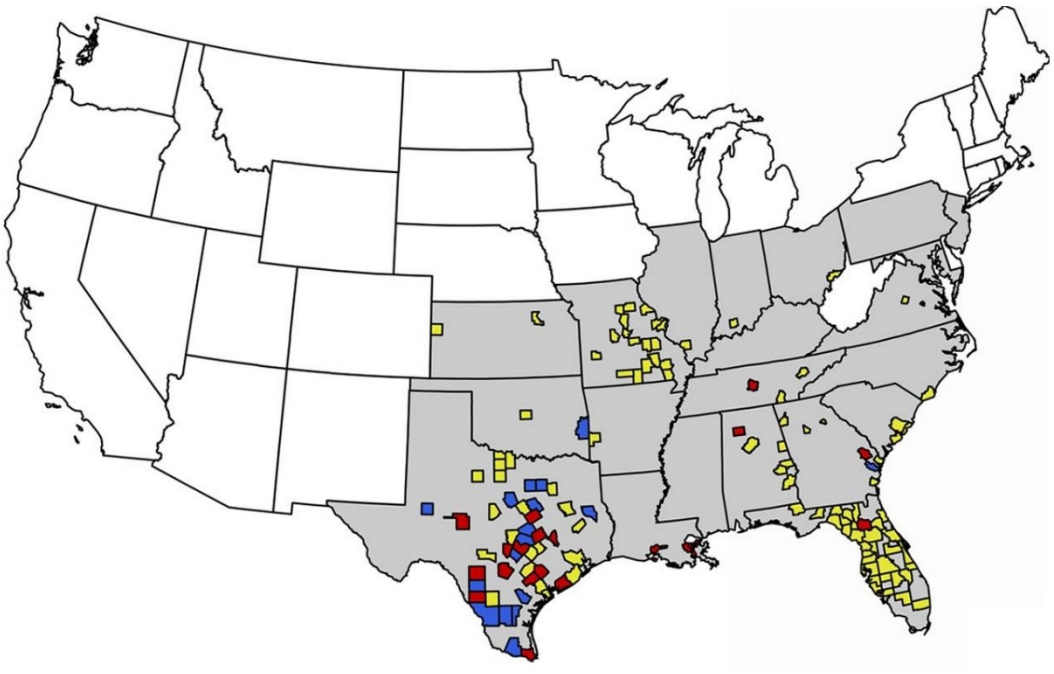


Figure 4 - Predicted distribution of *T. sanguisuga*. Here red counties denote *T. cruzi* infected *T. sanguisuga*, blue counties denote *T. cruzi* uninfected samples, and yellow counties denote the presence of *T. sanguisuga* on which no testing for *T. cruzi* was performed [25].

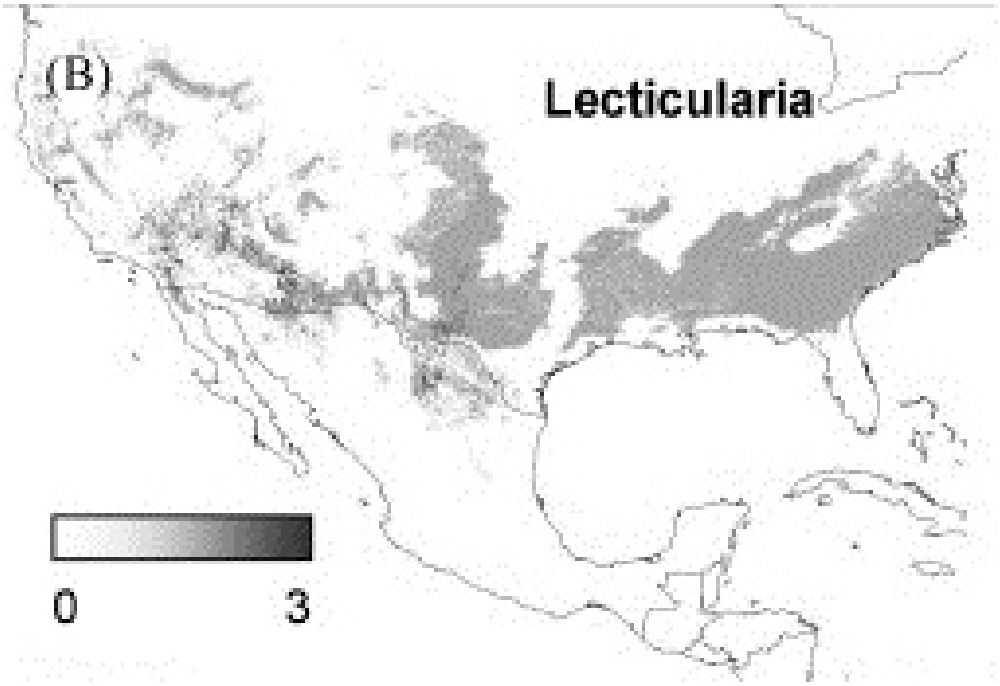


Figure 5 - Predicted distribution of the *Triatoma lecticularia* species complex (which includes *T. sanguisuga*). This model was generated using 18 unique collection points for *T. sanguisuga* in addition to points from other members of the species complex [4].

To determine the ecological niche of *T. sanguisuga* the present research undertakes a two-part approach. The first aspect is species distribution modeling of *T. sanguisuga* based on personal and publicly available collection records. The second aspect seeks to generate species distribution models for the reported associated animals of *T. sanguisuga* and compare the degree to which the distribution of each associated animal agrees with the predicted distribution of the vector. This will advance the body of knowledge on this topic by providing environmental variables that are predictive of *T. sanguisuga* presence and providing a new metric that may be used to narrow the list of animals associated with this species.

Chapter III

Specific Aims and Hypotheses

The present research aims to answer the questions: how do populations of *T. sanguisuga* and its associated parasite *T. cruzi* move through space? is the eastern woodrat the primary associate animal of *T. sanguisuga*? To answer these questions, the following hypotheses were formulated and tested:

Specific aim 1 – Populations of *T. sanguisuga* in southeastern Louisiana exist in focal, genetically related groups which leads to diverse genetic subgroups that are associated with particular discrete typing units of *T. cruzi*. The genetic homogeneity of *T. sanguisuga* identified in previous reports is not correct because of the limited dispersal range of the species.

Specific aim 2 – The most common bloodmeal source of *T. sanguisuga* in southeastern Louisiana is the eastern woodrat, *Neotoma floridana*, and human bloodmeals are rarely detected. Previous studies of the common bloodmeal sources of *T. sanguisuga* are incorrect because they over-represent humans.

Specific aim 3 – The ecological niche of the vector is tightly linked to the suitable habitat of small mammals that provide bloodmeals. Of the associated animals of *T. sanguisuga*, the suitable habitat of *N. floridana* is most highly predictive of the presence of *T. sanguisuga*.

Chapter IV

Methods and Materials

T. sanguisuga collections

Site 1

The majority of *T. sanguisuga* samples were collected in rural Orleans Parish at the home of the sixth autochthonous case of Chagas disease in the United States (termed "Site 1") [52]. Since the identification of this likely autochthonous case, the Wesson lab has been involved in weekly seasonal checks of the home and surrounding outbuildings for *T. sanguisuga*. Once each week from late Spring through early Fall, trained laboratory staff performed a thirty-minute search of the buildings on the property following a standard protocol. Data collected with each bug included: sex; life stage; specific location; whether the bug was alive, dead, or moribund; whether the location was indoors or outdoors; and the date of collection. These data were compiled in a Microsoft Excel database.

Other collection sites

Other *T. sanguisuga* specimens were collected from the areas adjacent to Site 1. These were also single-family, detached homes within the same contiguous section of bottomland hardwood forest. The remainder of the samples were collected haphazardly by laboratory staff or community members in the areas of Bush, Louisiana, Des Allemands, LA, Folsom, LA, Belle Chasse, LA, Bayou Gauche, LA, Morganza, LA, Bayou Sauvage (near New Orleans, LA), St. Gabriel, LA, New Iberia, LA, Poplarville, MS, Oxford, MS, and Blanco, TX (Figure 6).

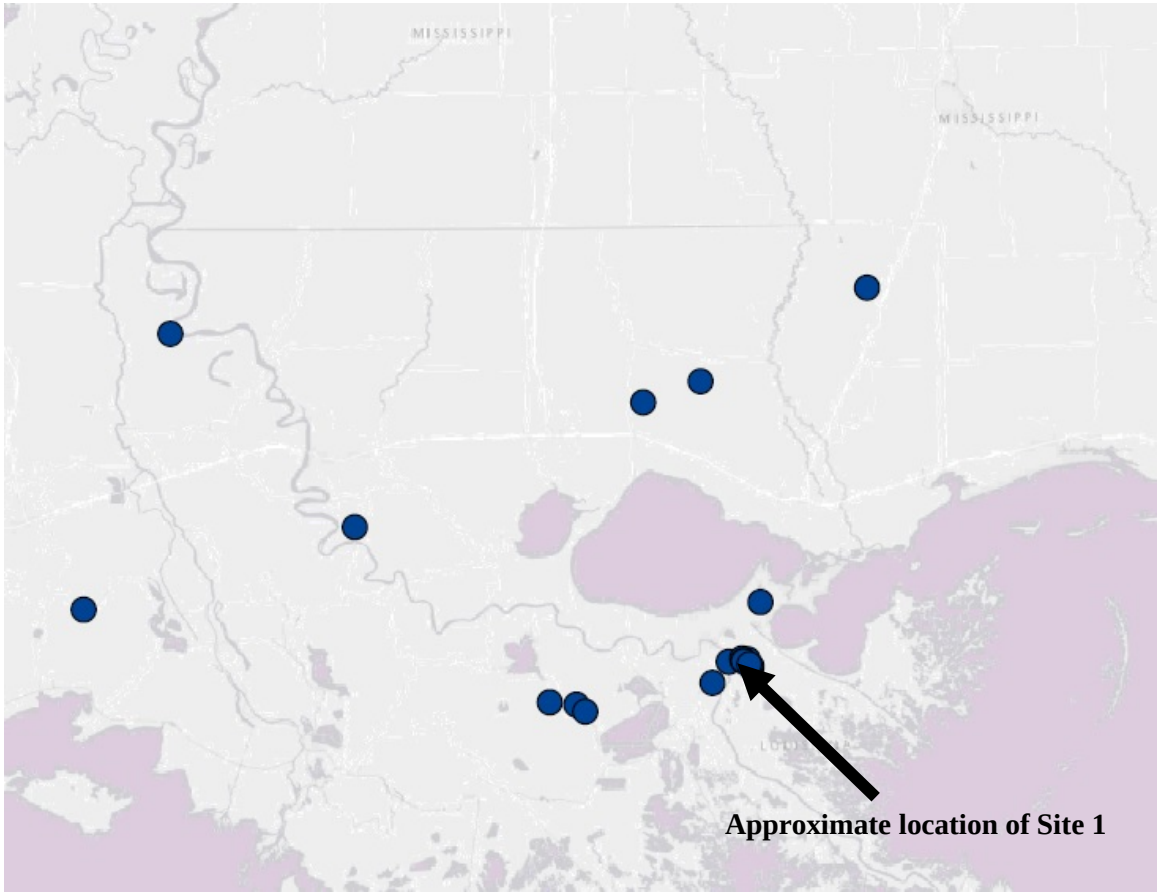


Figure 6 - Locations of *T. sanguisuga* samples collected haphazardly by laboratory staff or submitted to the laboratory by community members.

T. sanguisuga dissections

Dissections for *T. cruzi* PCR and bloodmeal analysis

T. sanguisuga that were collected alive or moribund were dissected as soon after collection as possible to recover living *T. cruzi.* Samples that were dead at the time of collection were either dissected the same day or stored at -20°C in 70% ethanol until dissections could take place. Dissections were performed on a clean benchtop using appropriate personal protective equipment to limit the possibility of self-inoculation with *T. cruzi*. A new disposable razor blade was used for each dissection. The posterior one-third of the abdomen was carefully removed and placed in a sterile 1.5 mL microcentrifuge tube. The dissected tissues were stored without buffer at -20°C until DNA extraction.

Dissections for *T. sanguisuga* population genetics

Previously dissected and frozen *T. sanguisuga* samples were retrieved, thawed, and dissected according to previously published methods with minor modifications [53]. Briefly, the legs of each sample were removed using flame-sterilized forceps and placed in a sterile 1.5 mL microcentrifuge tube. Only legs connected to the thorax were used to ensure a single DNA source. In the event that a specimen did not have legs the head or thorax, respectively, was used. In the event that only an abdomen remained, the sample was censored to avoid possible PCR inhibitors in the anterior abdomen [52].

DNA extraction

All samples were extracted using DNAzol (Molecular Research Center). The standard protocol was modified to enhance DNA recovery from desiccated and chitinous samples. Briefly, the samples were homogenized by sterile mortar and pestle or BeadBug

bead mill (Benchmark Life Sciences) using 0.25-gram plastic beads. The homogenized sample was incubated overnight at 55°C before continuing the extraction. Additionally, a pre-centrifuge step of 1,000 x gravity for 3 minutes was added to remove any large pieces of chitinous material before completing the manufacturer's protocol. All precipitated DNA samples were re-suspended in nuclease free water and stored at -20°C until analysis.

Conventional PCR

Conventional PCR was performed on samples of DNA from abdominal dissections and genomic DNA of *T. sanguisuga.* The various primer sets, their respective targets, and references are summarized in Table 4. The thermal cycling conditions and final primer concentration used for each set of primers are summarized in Table 5. All conventional PCR was performed with Apex Hot-Start MasterMix – Blue (Genesee Scientific) in total reaction volumes of 10 μL or 12.5 μL. Negative and positive amplification controls were included in every assay. Reaction products were separated by agarose gel electrophoresis in 1 x TAE and stained with ethidium bromide. Fragments were visualized by ultraviolet light transillumination. All samples electrophoresed with GeneGauge®100bp DNA Ladder H3(Caisson Labs). All conventional PCR was performed using a MyCycler thermal cycler system (Bio-Rad)

Table 4 - Primer sets utilized in the present research.

Primer Name	Sequence (5' – 3')	Target Species	Target locus	Fragment Size (base pairs)	Reference
TScytbF TScytbR	TAT GGG AGG CCC CGA TTC TA TGG GAG GCC TTG AAA TTT TCT	*T. sanguisuga*	Cytochrome b	218	*[a]
D71 D72	AAG GTG CGT CGA CAG TGT GG TTT TCA GAA TGG CCG AAC AGT	*T. cruzi*	24Sα rDNA	110, 120, 125	[54][b]
TC TC1 TC2	CCC CCC TCC CAG GCC ACA CTG GTG TCC GCC ACC TCC TTC GGG CC CCT GCA GGC ACA CGT GTG TGT G	*T. cruzi*	Spliced leader	0, 300, 350	[54][c]
V1 V2	CAA GCG GCT GGG TGG TTA TTC CA TTG AGG GAA GGC ATG ACA CAT GT	*T. cruzi*	18S rDNA	0, 155, 165, 175	[54][d]
L14816 H15173	CCA TCC AAC ATC TCA GCA TGA TGA AA CCC CTC AGA ATG ATA TTT GTC CTC A	Universal vertebrate	Cytochrome b	358	[55][e]

*Developed in the course of work (a) Primer set designed and used for investigating the small- and large-scale genetic relationships of collected *T. sanguisuga* (b, c, d) Primer sets used to determine the discrete typing unit (DTU) of *T. cruzi* DNA present in each sample (e) Primer set used to identify bloodmeal sources from *T. sanguisuga* specimens.

Table 5 - Conventional PCR cycling conditions utilized.

Primer Set	Primer Concentration	Initial Denature	Denature	Anneal	Extend	Number of Cycles	Final Extension	Hold
TScytbF TScytbR	0.4 µM	95°C 15 minutes	95°C 1 minute	51°C 1.5 minutes	72°C 1.5 minutes	35	N/A	4°C ∞
D71 D72	0.2 µM	95°C 15 minutes	94°C 1 minute	60°C 1 minute	72°C 1 minute	30	72°C 5 minutes	4°C ∞
TC TC1 TC2	0.2 µM	95°C 15 minutes	94°C 30 seconds	55°C 30 seconds	72°C 30 seconds	27	72°C 5 minutes	4°C ∞
V1 V2	0.2 µM	95°C 15 minutes	94°C 1 minute	50°C 1 minute	72°C 1 minute	30	N/A	4°C ∞
L14816 H15173	0.5 µM	95°C 15 minutes	95°C 30 seconds	60°C 30 seconds	72°C 30 seconds	35	N/A	4°C ∞

High-resolution melting analysis (HRMA)

HRMA technology

HRMA is a standard molecular technique that is capable of detecting single nucleotide polymorphisms (SNP) in a targeted sequence. As the name suggests, the analysis is based on capturing the melt curve of a real-time PCR assay in high-resolution. By measuring fluorescence changes in very small increments (≤ 0.2°C), the software is able to discriminate among sequences of the same length with different base-pair compositions. This change in fluorescence relative to temperature is a physical property of a given sequence because of the energy required to "melt" or separate each base pair is a function of the bases that are paired [56]. The Precision Melt Analysis software (Bio-Rad) used for these experiments allows the user to adjust two values to control how the software clusters the melt curves. The Tm difference threshold is the maximum allowable difference in melting temperature (Tm) allowed to consider two sequences the same. Bio-Rad suggests the default value of 0.15°C for routine applications. The second value that can be changed is the curve similarity setting. This setting can take any integer value from 1- 100, where 1 is most dissimilar and 100 is most similar. While Bio-Rad does not disclose how this metric is calculated, it suggests using lower values to cluster more heterozygous sequences, and higher values to cluster more homozygous sequences [57].

Primer design

A previously published primer set for the amplification of a section of the cytochrome b gene of *T. sanguisuga* [26] produced fragments longer than 600 base pairs. These fragments were determined to be too long to produce reliable HRMA results. The use of HRMA technology necessitated designing a novel primer set that amplifies a small

(<200 base pairs) variable region of the cytochrome b gene of *T. sanguisuga.* A previously suggested locus for which sequence data exists [26] was targeted for primer development. A representative sequence (HQ141280.1) was retrieved from the NCBI database and submitted to the NCBI/Primer-BLAST program using the default parameters. The primers were assessed manually for their coverage of sites known to be variable based on previous research [26], their Tm values, and length of produced fragment. A primer set, termed TScytbF/TScytbR (Table 4), was chosen, and reaction conditions were optimized using gradient PCR and validated using extracted *T. sanguisuga* genomic DNA and assorted mammalian DNA positive controls.

Thermal cycler and software

HRMA was performed using a single CFX96 Touch real- time PCR detection system (Bio-Rad) using Precision Melt Supermix (Bio-Rad). Machine calibration was performed prior to performing any HRMA experiments. Amplification success was determined based on log-scale amplification curves and negative first derivative curves generated in the CFX Manager software (Bio-Rad). HRMA analysis was carried out in Precision Melt Analysis software (Bio-Rad). The default settings were modified depending on the samples being analyzed. For bloodmeal analysis, curve similarity was set to 80 with a Tm difference threshold of 0.08°C. For population genetics studies of *T. sanguisuga* and *T. cruzi*, the most conservative settings of the software were applied (curve similarity = 100, Tm difference threshold = 0.05°C). Optimization of the HRMA settings for bloodmeal analysis performed using various positive control DNA samples of vertebrates. All reactions were performed in a final volume of 10 μL. Samples that failed to amplify or produced irregular curves were repeated for confirmation. Samples that

failed to amplify or repeatedly showed unsatisfactory melting curves were removed from analysis. Positive and negative controls were included in every HRMA assay. In all cases, representative sequences of each cluster called by the software were sent to Eton Bioscience, Inc. (Durham, NC) for sequence determination. Sequences were viewed and trimmed using 4Peaks v1.8 (Nucleobytes), combined into contigs using MEGA 7 [58], and confirmed using NCBI's nucleotide BLAST.

Primers and thermal cycling

Thermal cycling conditions required slight modification when using the Precision Melt Supermix. Modified thermal cycling conditions as well as melt curve acquisition parameters are summarized in Table 6.

Table 6 - HRMA cycling conditions utilized.

Primer Set	Primer Concentration	Initial Denature	Denature	Anneal	Extend	Number of Cycles	Dissociation	Duplex Formation	Melt Curve
L14816 H15173	0.5 µM	95°C 2 minutes	95°C 30 seconds	60°C 30 seconds	72°C 30 seconds	35	95°C 30 seconds	65°C 1 minute	65°C - 95°C 0.2°C/step 10 second dwell
TScytbF TScytbR	0.4 µM	95°C 2 minutes	95°C 1 minute	51°C 1.5 minutes	72°C 1.5 minutes	35	95°C 30 seconds	65°C 1 minute	65°C - 95°C 0.2°C/step 10 second dwell

HRMA curves were constructed from 10 second image acquisitions in 0.2°C increments by the Precision Melt Analysis software.

Haplotyping of human DNA samples from *T. sanguisuga*

Discrimination of human cytochrome b HRMA clusters

Samples that tested positive for human DNA during bloodmeal analysis were further analyzed to determine whether the human DNA was attributable to individuals who had worked on the project. Discrimination amongst sequences was accomplished by highly conservative HRMA curve analysis of the cytochrome b sequences generated during bloodmeal analysis. The settings of the software were adjusted to a curve similarity of 100 with a Tm difference threshold of 0.05°C. The present research was not approved for the generation of sequence level data of human samples, so HRMA clusters were used to approximate the number of different cytochrome b sequences observed.

Human control samples

Available staff who took part in the handling and DNA extraction of the *T. sanguisuga* abdominal samples were contacted regarding participation in the research project. Willing individuals (n=11) were asked to submit an oral rinse sample using nuclease free water. Collected samples were extracted for DNA no more than 36 hours after collection. Buccal cells were pelleted from the sample and the DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen). DNA samples were eluted into water and stored at -20°C until analysis. The present research was approved by the institutional review board of Tulane University - #16-892815.

Geospatial distribution of *T. sanguisuga* and *T. cruzi* genotypes

Large-scale spatial distribution of *T. sanguisuga* genotypes

Laboratory collected *T. sanguisuga* specimens with reliable collection data were consolidated and assigned random, computer-generated identification numbers. From

each unique collection site (or home), a maximum of five samples were selected by taking the lowest five values of the sorted and coded list. These samples were analyzed by HRMA using the previously mentioned cycling parameters. Precision Melt Analysis software (Bio-Rad) was used to analyze the HRMA data and call clusters. The software settings were adjusted to the most stringent settings for designating a cluster: (Curve sensitivity= 100; Tm difference threshold = 0.05°C).

A representative sample from each cluster was re-amplified in a 25µL conventional PCR reaction. Amplification was confirmed by agarose gel electrophoresis of 5µL of the product, and the remaining 20µL were sent to Eton Bioscience for cleaning and sequencing. Sanger sequencing was performed on an Applied Biosystems 3730xl DNA Sequencer using BigDye® Terminator v3.1 (Applied Biosystems). Trace files were evaluated and trimmed with 4Peaks v3.1 (Nucleobytes) and contigs were constructed in MEGA 7 [58] (Full sequences are listed in Appendix A.)

The cytochrome b sequences of these samples were analyzed following previously published methods [26]. Briefly, genotypes were statistically analyzed for correlation with sex, *T. cruzi* infection status, and month and year of collection using GraphPad Prism 7 software (GraphPad Software, La Jolla CA, USA). Phylogenetic trees and analysis were completed in MEGA 7 [58] using maximum likelihood and neighbor-joining methods. Genetic analysis was performed in DNAsp v5.10.1 [59]. A median-joining network analysis [60] was conducted using Network (fluxus-engineering.com) with post-processing to simplify the network, as previously published [26]. Additionally, these collection data were plotted in ArcMap 10.4 for Desktop (ESRI) to visualize the distribution of haplotypes across large distances.

Small-scale spatial distribution of *T. sanguisuga genotypes*

To determine that applicability of the cytochrome b locus to small-scale spatial relationships, the genomic DNA samples from Site 1 were used. All samples were analyzed by HRMA as described previously. Precision Melt Analysis software (Bio-Rad) was used to analyze the HRMA data and call clusters. The software settings were adjusted to the most stringent settings for designating a cluster: (Curve shape sensitivity = 100, Tm difference threshold = 0.05°C).

A representative sample from each cluster was re-amplified in a 25µL conventional PCR reaction. Amplification was assessed by gel electrophoresis of 5µL of the product. The remaining 20µL were sent to Eton Bioscience for cleaning and sequencing. Sanger sequencing was performed on an Applied Biosystems 3730xl DNA Sequencer using BigDye® Terminator v3.1 (Applied Biosystems). Trace files were evaluated and trimmed with 4Peaks v3.1 (Nucleobytes). (Full sequences are listed in Appendix A.)

These samples were further geocoded using decimal degree measurements to their precise collection location in ArcMap 10.4. As some of the buildings at Site 1 have two stories, the elevation of the collection was recorded as 1 or 15 to reflect the approximate distance (in feet) of each collection above the ground. The points were projected using the 2011 South Louisiana state projection (feet) in ArcMap. Genetic, phylogenetic, and network analyses were performed on these data following the methods used for the large-scale spatial data. The geocoded coordinates (x, y, and z) were analyzed by optimized spatial clustering and hot-spot algorithms using year and cytochrome b haplotype as clustering variables. All spatial analysis was performed in ArcMap 10.4.

Large- and small-scale spatial distribution of *T. cruzi* DTUs

Abdominal DNA samples were tested by conventional PCR using the specified *T. cruzi* primers in Table 4. Samples determined to be positive were analyzed by HRMA with the V1/V2 primer set. A representative sample from each cluster called by the software was submitted to Eton Bioscience for sequencing. Sequencing, alignment, and construction of contigs were performed as before. To determine the spatial distribution of various *T. cruzi* DTUs, geocoded samples with defined *T. cruzi* genotypes were plotted in ArcMap 10.4 for visualization.

Ecological niche modeling of *T. sanguisuga*

T. sanguisuga collection records database construction

Presence data of *T. sanguisuga* were compiled from samples submitted to the laboratory by community members and publicly available records. Community submitted samples were received over several years from concerned community members for positive identification by laboratory staff. publicly available collection records were retrieved from the Louisiana State Arthropod Museum [61], iDigBio [62], and iNaturalist [63]. Geocoded collection data were collected and cross-referenced with the Symbiota Collections of Arthropods Network (SCAN) database [64]. Data were cleaned to remove samples with obviously incorrect coordinates (e.g. in a body of water). Only presence records containing latitude, longitude, and an acknowledged coordinate error of less than 3,000 meters were included.

Maximum entropy model of *T. sanguisuga*-suitable habitats

The database of *T. sanguisuga* collection records was cleaned to contain only geographically unique occurrences. The final environmental layers used are summarized in Table 7. To create the leaf area index variable (LAI) in Table 7, monthly measurements from January 2001 – December 2010 were collected and summarized using the maximum, minimum, and mean cell statistics in ArcMap 10.4. All variables were trimmed to the boundary of the contiguous United States using census data [65]. MaxEnt v3.3.3k [66] was used to create a maximum entropy (MaxEnt) model of habitats suitable for *T. sanguisuga.* Presence data was loaded from the previously constructed database. Raw output was generated and averaged by the software from 50 cross-validated models. Jackknife resampling for variable importance was performed, and

multidimensional environmental similarity surface (MESS) was used during projection of the models. A coordinate bias file (CBF) was used for the contiguous United States to control for the wide range of latitudes from which specimens were observed. The CBF was generated using SDMtoolbox v1.1c [67] in ArcMap software. The "Urban and Built" layer was removed from the AVHRR land cover raster *a priori* as initial models suggested a confounding effect.

Associated animal distributions and calculations

MaxEnt models were created for the associated mammals listed in Table 2 using the previously mentioned parameters with few exceptions. Occurrence data were collected from sources summarized in Table 8. All species were modeled using 20 cross-validated models. The previously-generated CBF was utilized to control for coordinate bias within these models.

The difference of each model was calculated pair-wise against the MaxEnt model for *T. sanguisuga* and ranked from least to most dissimilar using ENMtools v1.4.4 [68]. Additionally, the rasterized MaxEnt models for each of the associated animals were added as environmental layers in a MaxEnt model of *T. sanguisuga* occurrence points to assess the association amongst the presence of *T. sanguisuga* and the suitable habitats for the associated animals.

Table 7 - Sources for habitat modeling of *T. sanguisuga* and its associated animals.

Dataset Name	Description	Data Source	Layers		Reference
WorldClim v1.4	Bio-climactic variables aggregated from 1960-1990. 30 second arcs.	Multiple major climate databases. All data is a summary of >10 years of data points	Annual mean temperature Mean diurnal range Isothermality Temperature seasonality Maximum temperature of warmest period Minimum temperature of coldest period Annual temperature range Mean temperature of wettest quarter Mean temperature of driest quarter Mean temperature of warmest quarter	Mean temperature of coldest quarter Annual precipitation Precipitation of wettest period Precipitation of driest period Precipitation seasonality Precipitation of wettest quarter Precipitation of driest quarter Precipitation of warmest quarter Precipitation of coldest quarter	[66]
Leaf Area Index	Area covered by leaves. Composites of monthly samples from 2001-2010.	MODIS* reflectance measurement	Leaf area index: 1 - 100		[67, 68]
Land Cover	Classification of land into 12 categories	MODIS* derived classifications	Evergreen needleleaf forest Evergreen broadleaf forest Deciduous needleleaf forest Deciduous broadleaf forest Mixed forest Woodland Wooded grassland	Closed shrubland Open shrubland Grassland Cropland Bare ground Urban and Built*	[69, 70]

The Urban and Built layer was deleted from the raster before construction of the final models because it appeared to be highly correlated with *T. sanguisuga* presence due to sampling bias.

Table 8 - Resources used for presence data of associated animals of *T. sanguisuga.*

Associated animal	**Scientific name**	**Collection Data Source**	**Reference**
Eastern woodrat	*Neotoma floridana*	Integrated Digitized Biocollections	[62]
Raccoon	*Procyon lotor*	Integrated Digitized Biocollections	[62]
Squirrel	*Sciurus carolinensis or Sciurus niger*	Integrated Digitized Biocollections	[62]
Striped skunk	*Mephitis mephitis*	Integrated Digitized Biocollections	[62]
Opossum	*Didelphis viriginiana*	Integrated Digitized Biocollections	[62]
Hispid Cotton Rat	*Sigmodon hispidus*	Integrated Digitized Biocollections	[62]
Cotton mouse	*Peromyscus gossypinus*	Integrated Digitized Biocollections	[62]
Armadillo	*Dasypus novemcinctus*	Integrated Digitized Biocollections	[62]

Chapter V

Aim 1 – Results and Discussion

Identification of genetically similar groups using HRMA

There currently exists little information on the spatial movements of *T. sanguisuga* adults. The frequency with which populations of genetically related bugs move and mix is of particular interest as it relates to the development of effective vector control strategies. To capture the breadth of *T. sanguisuga* population genetics within southeastern Louisiana, similar analyses were undertaken on both a large and small scale. The large-scale data set was composed of *T. sanguisuga* specimens from across southeastern Louisiana (Figure 6), while the home of the sixth case of autochthonous Chagas disease (deemed, Site 1) was utilized for the small-scale analysis.

Tissues were dissected from 561 *T. sanguisuga* specimens from across southeastern Louisiana. These samples were directly tested by HRMA using the primer set TScytbF/TScytbR (Table 4). Of the extracted samples, 177 were found suitable for further analysis based on satisfactory amplification and melting characteristics (Figure 7). These data were then analyzed at a large and small geospatial scale. Overall, ten distinct cytochrome b haplotypes were identified (Figure 8).

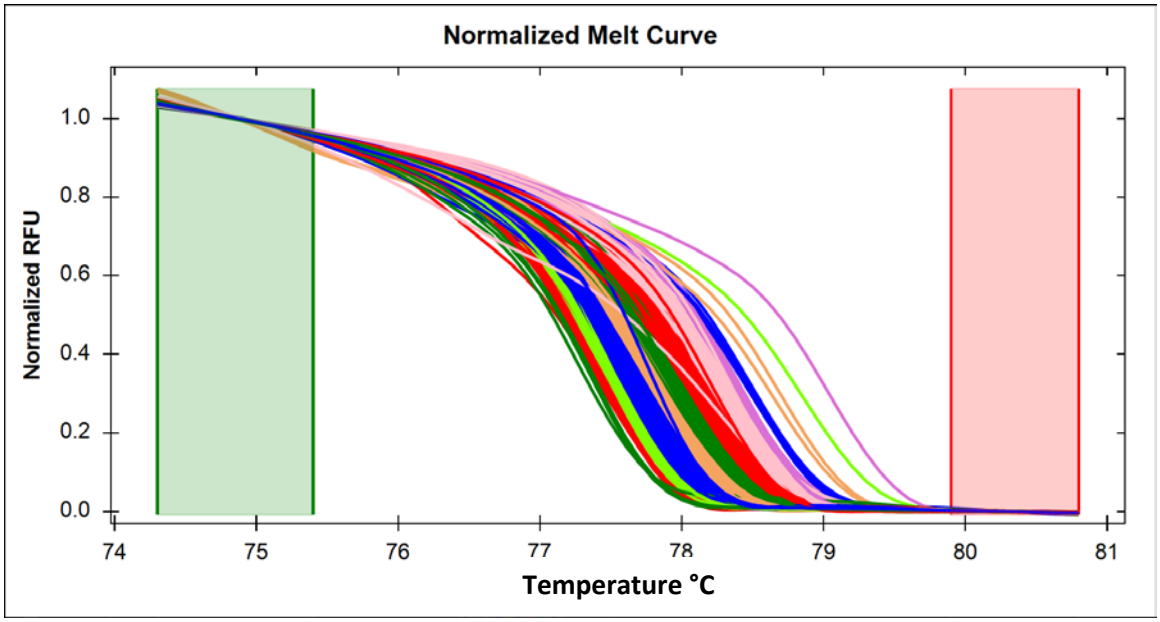


Figure 7 - Normalized melt curves from *T. sanguisuga* cytochrome b HRMA assay. Each color represents a distinct genetic sequence. The green and red regions flanking the curves identify the pre- and post-melt ranges used for normalization. Due to software limitations, after 10 clusters are identified the colors are re-used in the graphs.

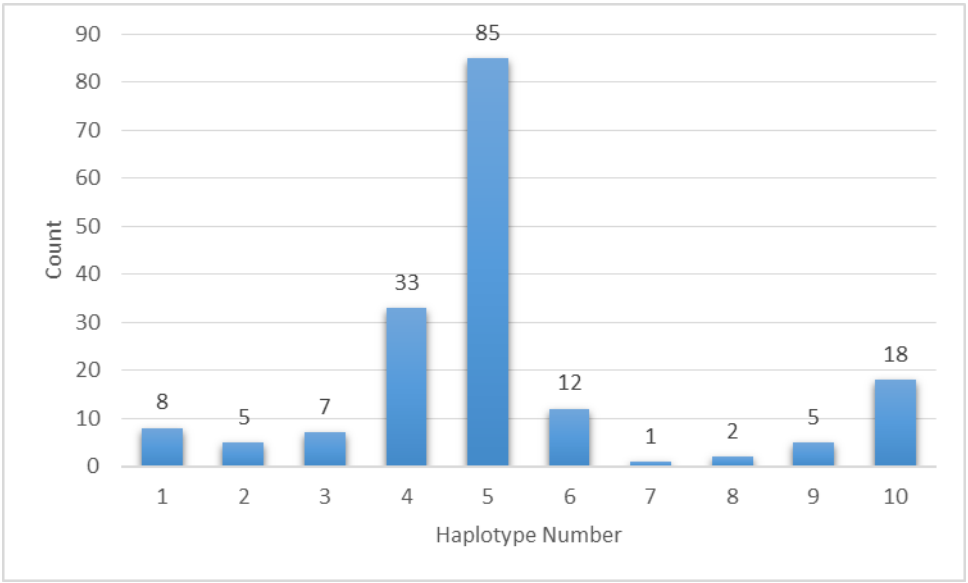


Figure 8 - Number of each cytochrome b genotype identified by HRMA

Large-scale population genetics

A random sub-sample was taken from each geographically unique location to determine the large-scale genetic differences of the cytochrome b gene in *T. sanguisuga* in southeastern Louisiana. Samples were assayed without duplication. Samples that failed to amplify or display smooth melt curves (n = 18) were not considered for analysis. A total of forty-one samples from across southeastern Louisiana were separated into eight haplotypes based on the amplified portion of the cytochrome b gene (Figure 9).

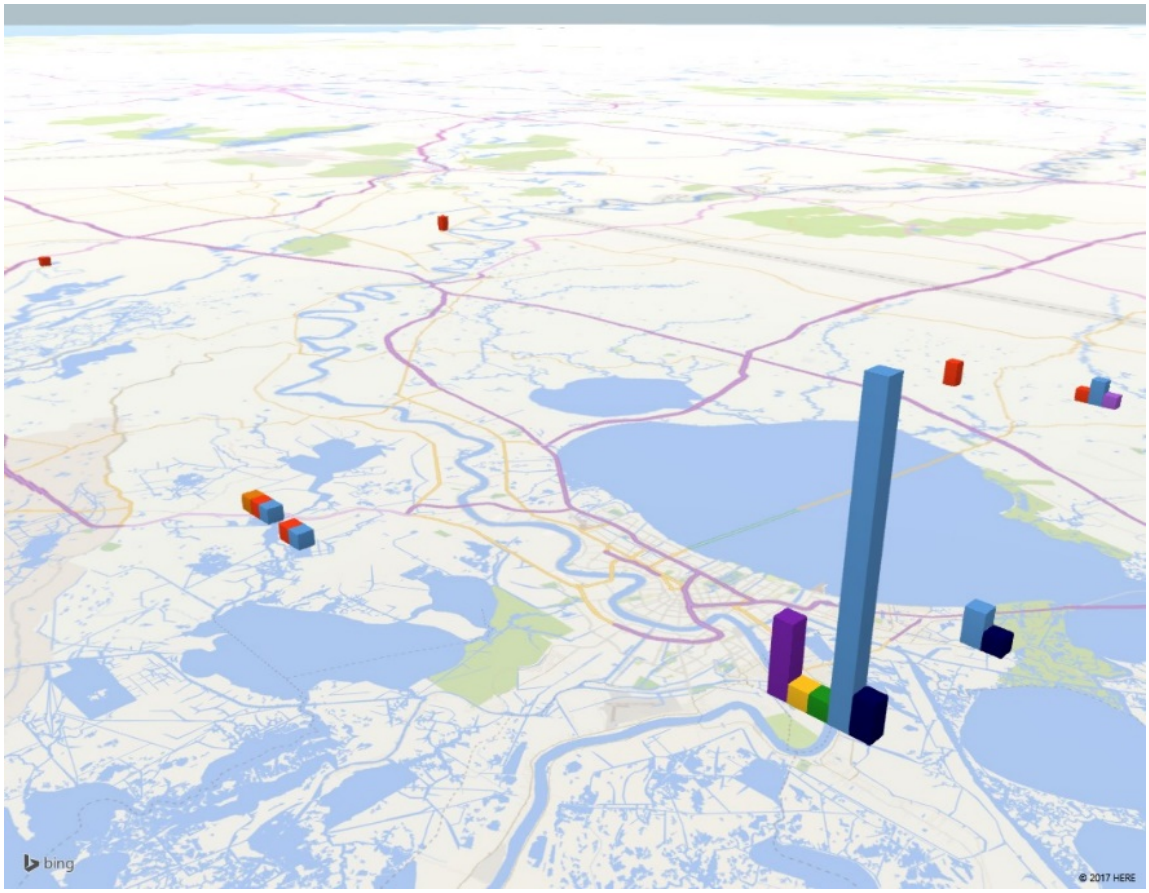


Figure 9 - Counts of haplotypes by location. Light blue = Haplotype 1, red = Haplotype 2, dark purple = Haplotype 3, dark blue = Haplotype 4, green = Haplotype 5, orange = Haplotype 6, yellow = Haplotype 7, light purple = Haplotype 8. Samples from rural Orleans Parish are aggregated for clarity.

Genetic analysis

Sequence alignment and construction of phylogenetic trees took place in MEGA 7 [58] using maximum likelihood (Figure 10) and neighbor-joining (Figure 11) methods. Both trees separated the eight haplotypes into three genetic clusters. Although the groupings are only supported by marginal bootstrap values, it is interesting that both methods of tree construction resulted in the same terminal groupings despite a disagreement, with poor bootstrap support, as to how the groups are related.

Genetic analysis was performed in DNAsp v5.10.1 [59]. The eight identified haplotypes were found to have a haplotype diversity of 0.673. The 207 base pair sequences were found to have 9 polymorphic sites. Three of these were singleton variable sites and 6 were parsimony informative. Fu's F statistic ($F = -1.225$, p-value >0.1) and Tajima's D statistic ($D = -0.52749$, p-value > 0.1), which test the genetic neutrality of a population, were calculated for this subset. The number of observed haplotypes was not larger or smaller than expected, which suggest that the population has not undergone a recent genetic expansion or bottleneck.

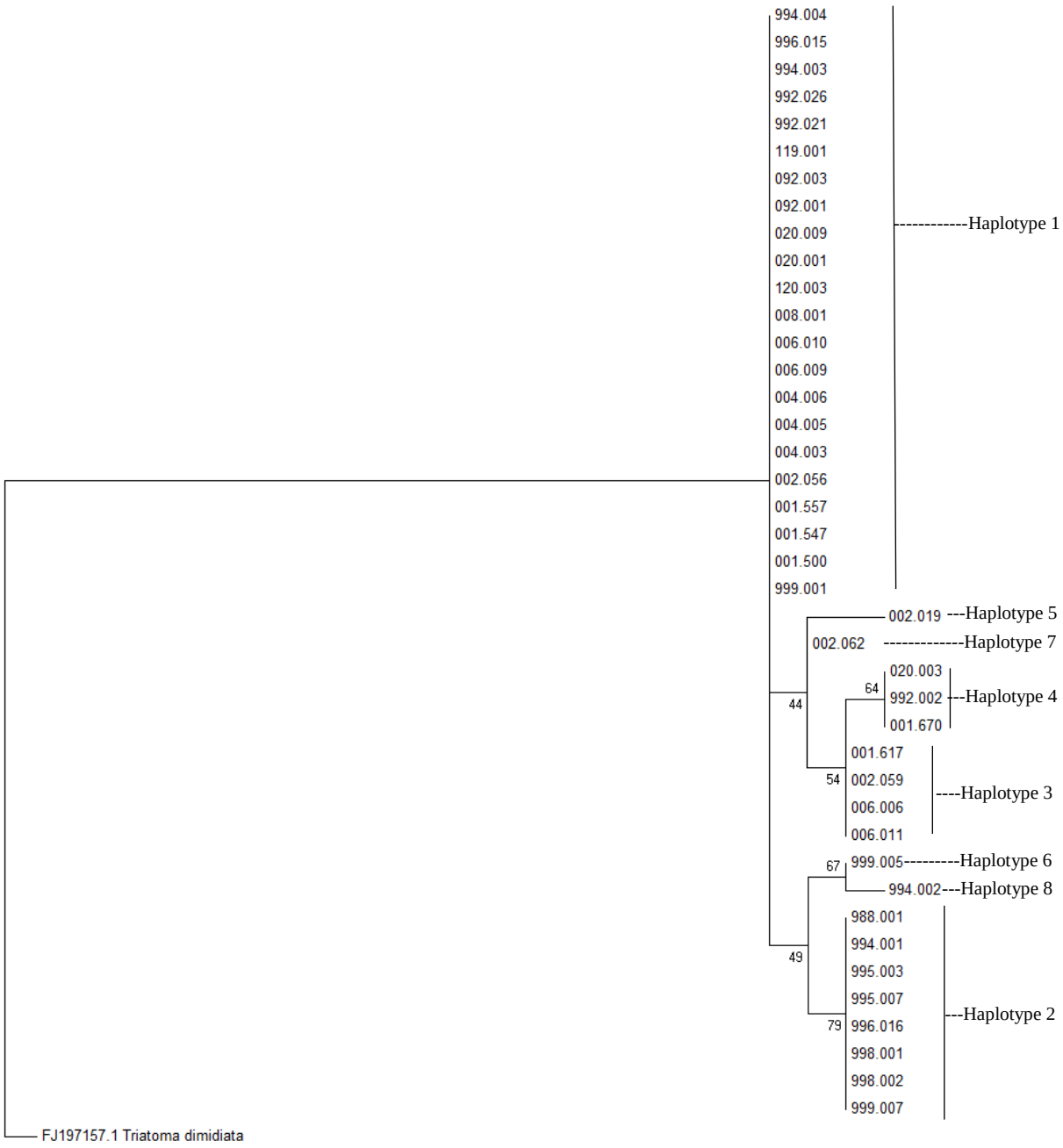


Figure 10 - Molecular Phylogenetic analysis by maximum likelihood method. The evolutionary history was inferred by using the maximum likelihood method based on the Tamura-Nei model [69]. The tree with the highest log likelihood (-409.0276) is shown. Bootstrap values (1,000 replicates) are shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying the neighbor join algorithm to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 42 nucleotide sequences. All codon positions were used. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA7 [58].

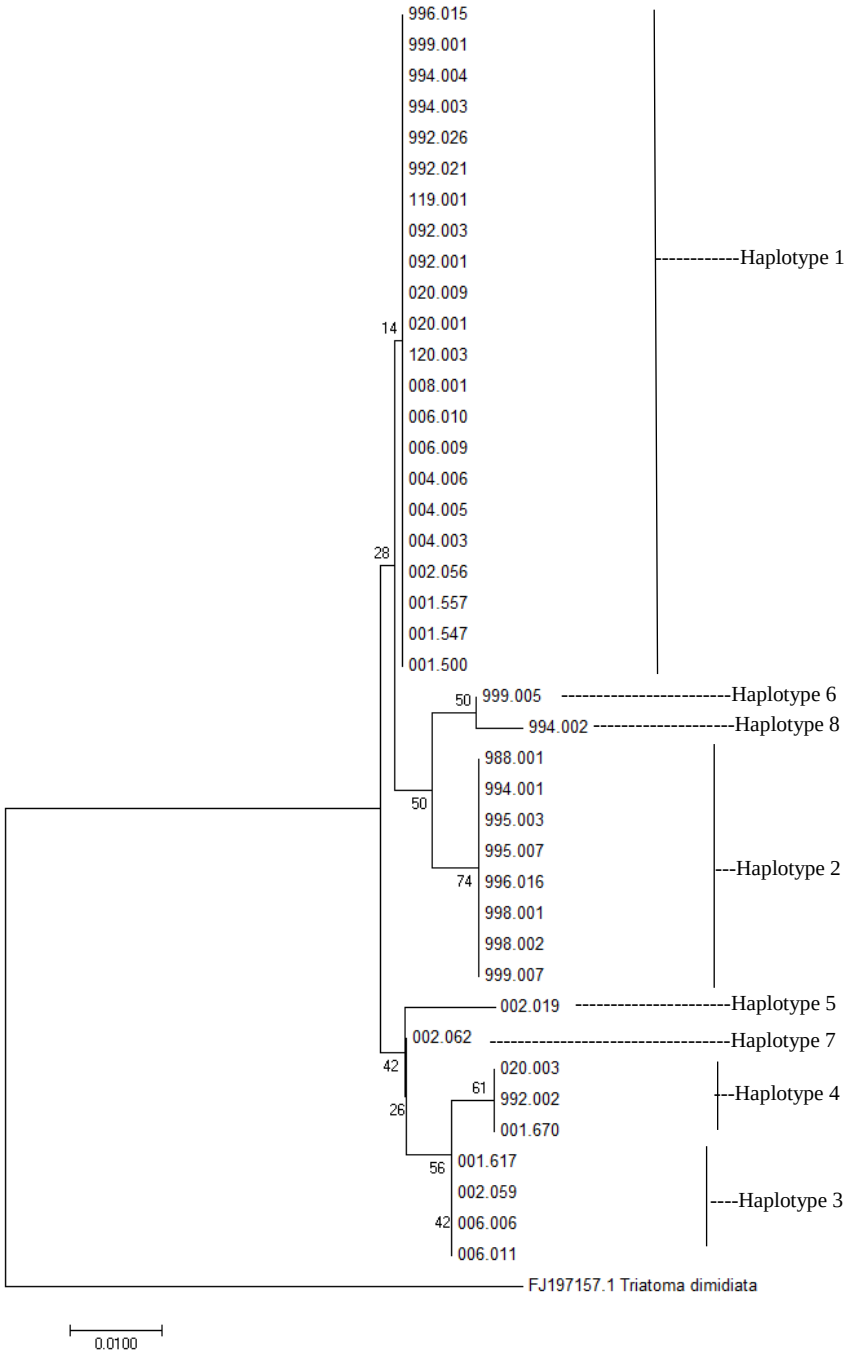


Figure 11 - Evolutionary relationships of taxa. The evolutionary history was inferred using the neighbor-joining method [100]. The optimal tree with the sum of branch length = 0.14083830 is shown. Bootstrap values (1000 replicates) are shown next to the branches [101]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Tamura-Nei method [58] and are in the units of the number of base substitutions per site. All codon positions were used. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA7 [58].

A median-joining network analysis [60] was conducted using Network (fluxus-engineering.com) with pre- and post-processing as previously published [26, 70]. The resulting network (Figure 12) identified the same three genetic groups identified by the maximum likelihood and neighbor-joining trees. At the large-scale, no relationship was found between geographic location and haplotype.

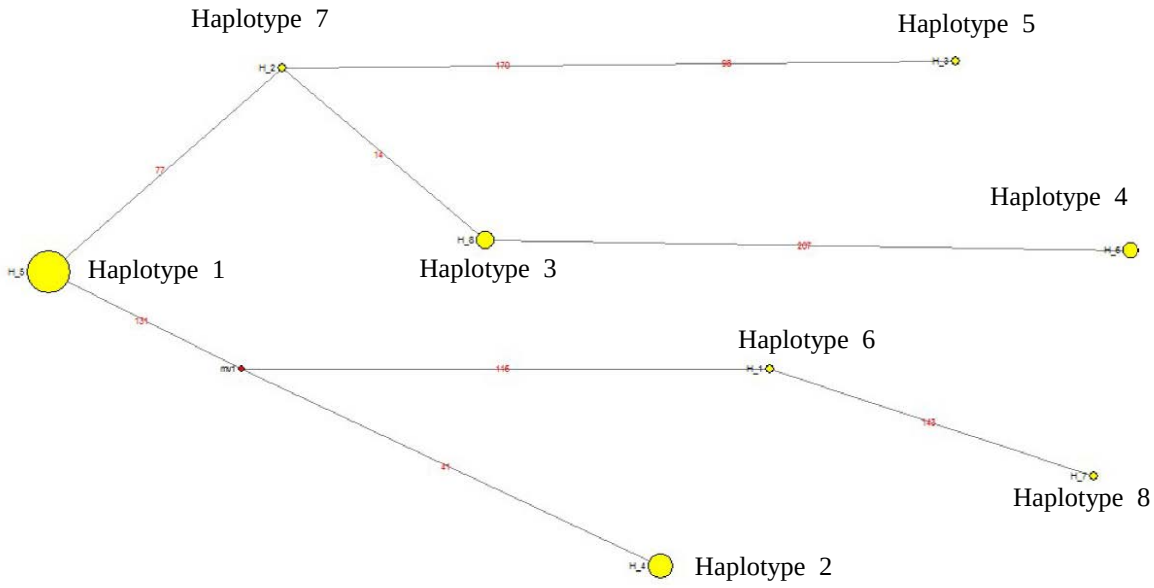


Figure 12 - Median-joining network of *T. sanguisuga* haplotypes. Default assumptions were utilized for network construction and standard post-processing was used to simplify the network. The network diagram displays the same topology as both phylogenetic trees.

These data combined with the data generated previously [26] suggest that the cytochrome b gene is not an appropriate genetic locus for large-scale population genetics studies of *T. sanguisuga*. Given the low dispersal range [5] and long reproductive interval [23] of *T. sanguisuga*, one would expect to find divergent populations over large distances. Although the cytochrome b gene is well characterized and is frequently used in population genetics studies of other insects [5], it does not provide satisfactory results in *T. sanguisuga.* A previous study of the population genetics of *T. sanguisuga* in southeastern Louisiana also found no apparent relationship between geographic location and cytochrome b haplotype [26]. Despite this shortcoming, this gene is used in the present research because the cytochrome b gene was previously found to yield more phylogenetic detail than the 16S locus [26]. Although it is not informative with respect to the large-scale population genetics of *T. sanguisuga*, this genetic locus suggests the possibility of sub-speciation of *T. sanguisuga* in Louisiana. De la Rua et al. noted an approximate 5% genetic distance between their two major genetic groups. Such genetic distance of the cytochrome b gene is higher than the average 2.8% cytochrome b genetic distance observed in *Triatoma rubida* [71], but less than the 8-9% genetic distance that has been used to demonstrate distinct species within the related *Triatoma brasiliensis* group [72] and Phyllosoma complex [71, 73]. Similar genetic grouping was observed in the present research, however no group approached a genetic distance of 5% (Figure 11, Figure 15). Though the analyzed locus was not able to identify the same divergent populations observed by De la Rua et al., the consistent grouping of the three genetic groups suggests that sub-populations, or even subspecies, remain a possibility. This is

important because it is possible that the different genetic groups possess differing vectoral capacities or behaviors that may increase human-bug interactions.

It is possible that the sequence variation observed in this population of insects is not sufficient to warrant a subspecies classification. Information from multiple genetic loci as well as interbreeding studies are necessary to confirm the level of divergence between the observed genetic groups.

The large-scale population genetics in the present research are limited by the length of the DNA fragment amplified for analysis. Although the use of HRMA allowed for the cost-effective analysis of samples, the technology is most effective with amplicons shorter than those analyzed by others [26]. In future studies, this limitation may be overcome using large platform technologies like 454 pyrosequencing, microsatellite analysis, or SNP detection from reduced representation libraries [74, 75].

Small-scale population genetics

All of the *T. sanguisuga* samples collected from Site 1 were utilized to determine the small-scale genetic differences at the home over time (Table 9). Samples were assayed without duplication. Samples that failed to amplify or display smooth melt curves were not considered for analysis (n = 154). A total of 164 samples were divided into ten haplotypes based on the amplified portion of the cytochrome b gene. For spatial analysis, the haplotype determination of each specimen was plotted in 3-dimensions on a map of Site 1[†] that was created using a combination of satellite imagery, architectural drawings, and detailed spatial data recorded at the time of each bug's collection (Figure 13).

[†] Due to the confidentiality of health data, all figures and spatial data for Site 1 are presented without a base map or other identifying information.

Table 9 - Summary of *T. sanguisuga* samples used for analysis.

Site Number	Total number of *T. sanguisuga extracted*	Number of samples suitable for analysis	Percentage of suitable samples
Site 1	318	164	52%

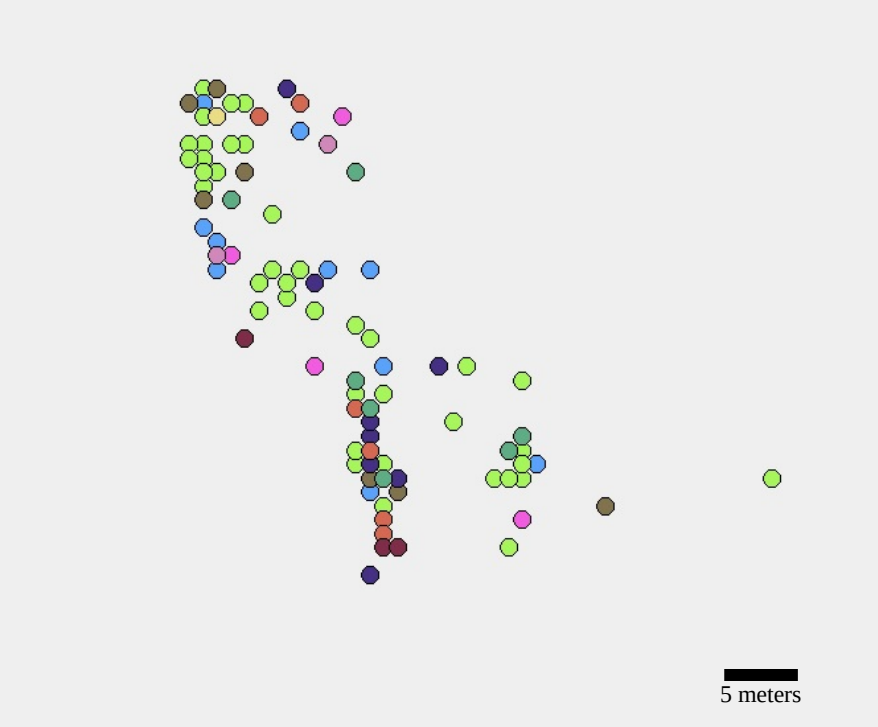


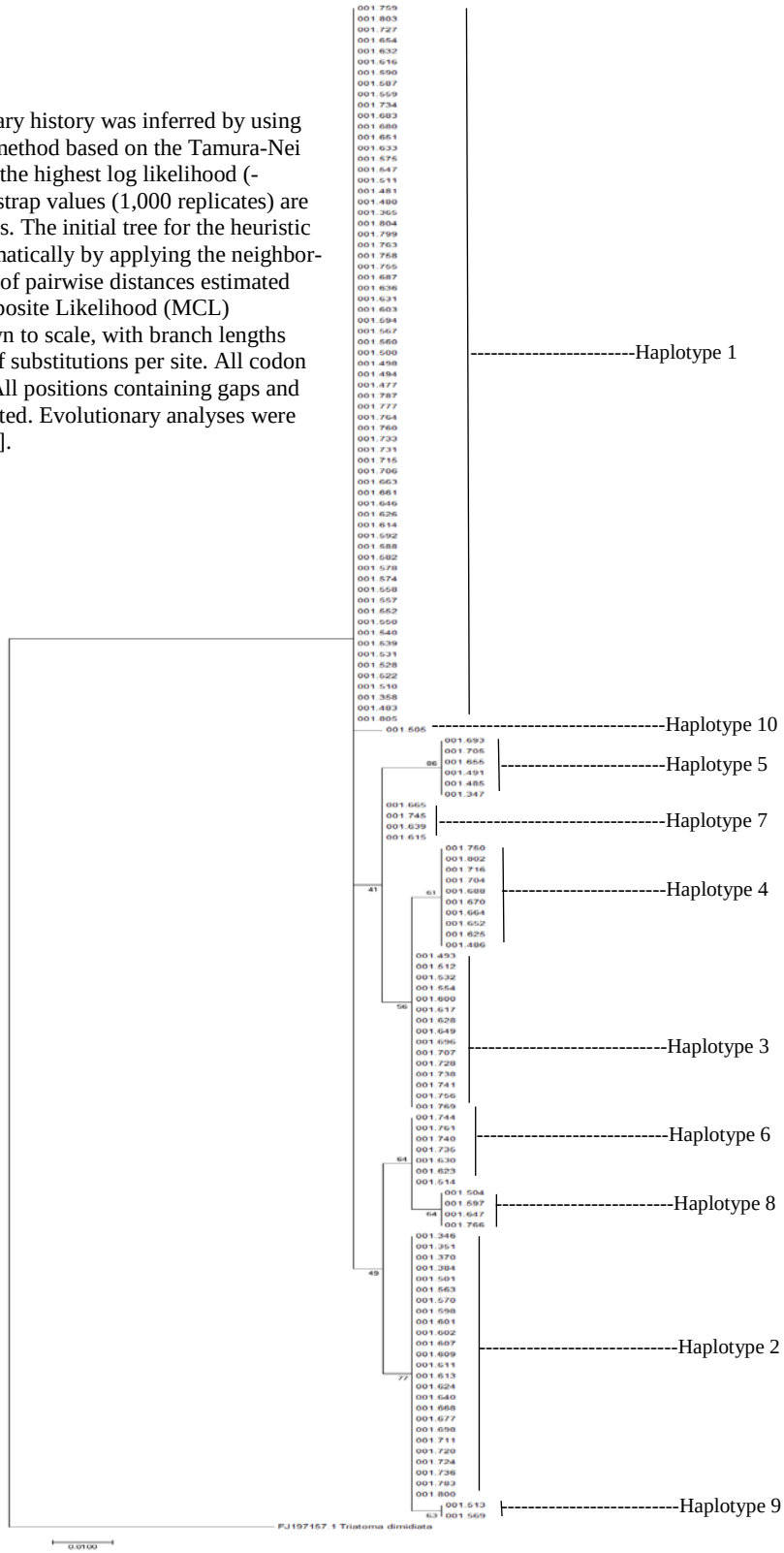
Figure 13 - Distribution of *T. sanguisuga* haplotypes at a single residence. Each color represents a distinct haplotype. Because the points were plotted in three-dimensions, only the top-most layer is displayed here. All calculations were carried out in three-dimensional space.

Genetic analysis

Sequence alignment and construction of phylogenetic trees took place in MEGA 7 [58] using maximum likelihood (Figure 14) and neighbor-joining (Figure 15) methods. Like the trees from the large-scale analysis, the trees calculated from the small-scale data separated the ten haplotypes into three distinct genetic clusters. As in the large-scale analysis, the bootstrap support for the trees is low. However, both analysis methods identified the same three genetic clusters.

Genetic analysis was performed in DNAsp [59]. The ten haplotypes identified were found to have a haplotype diversity of 0.725. The 207 base pair sequences were found to have 11 polymorphic sites. One of these was a singleton variable site and 10 were parsimony informative. Fu's F statistic ($F = -0.515$, p-value >0.1) and Tajima's D statistic ($D = -0.09271$, p-value > 0.1) were calculated as before and found to be insignificant. The number of observed haplotypes was not larger or smaller than expected, which suggest that the population has not undergone a recent genetic expansion or bottleneck.

Figure 14 - The evolutionary history was inferred by using the maximum likelihood method based on the Tamura-Nei model [69]. The tree with the highest log likelihood (-423.2044) is shown. Bootstrap values (1,000 replicates) are shown next to the branches. The initial tree for the heuristic search was obtained automatically by applying the neighbor-join algorithm to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All codon positions were included. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA7 [58].



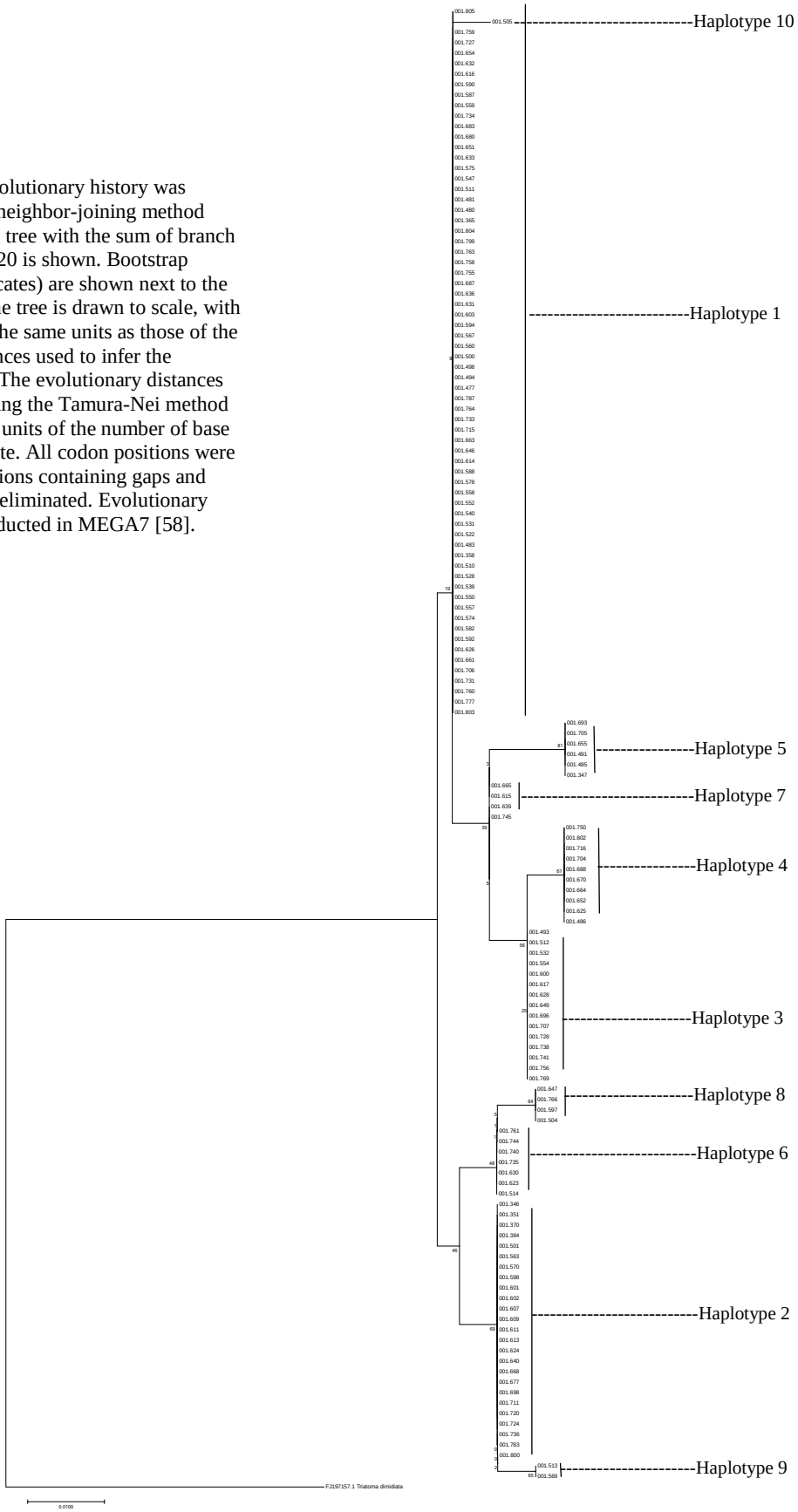


Figure 15 - The evolutionary history was inferred using the neighbor-joining method [100]. The optimal tree with the sum of branch length = 0.15180620 is shown. Bootstrap values (1000 replicates) are shown next to the branches [101]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Tamura-Nei method [69] and are in the units of the number of base substitutions per site. All codon positions were included. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA7 [58].

A median-joining network analysis [60] was conducted using Network (fluxus-engineering.com) with pre- and post-processing as previously published [26, 70]. The calculated network (Figure 16) identified the same three terminal genetic groupings identified by the maximum likelihood and neighbor-joining analyses at both the large and small geospatial scales. Median-joining network analysis was used in place of maximum parsimony analysis. As both methods utilize Farris's heuristic algorithm for adding vertices, the results are equivocal with the exception that median-joining networks do not settle ties during calculation [76].

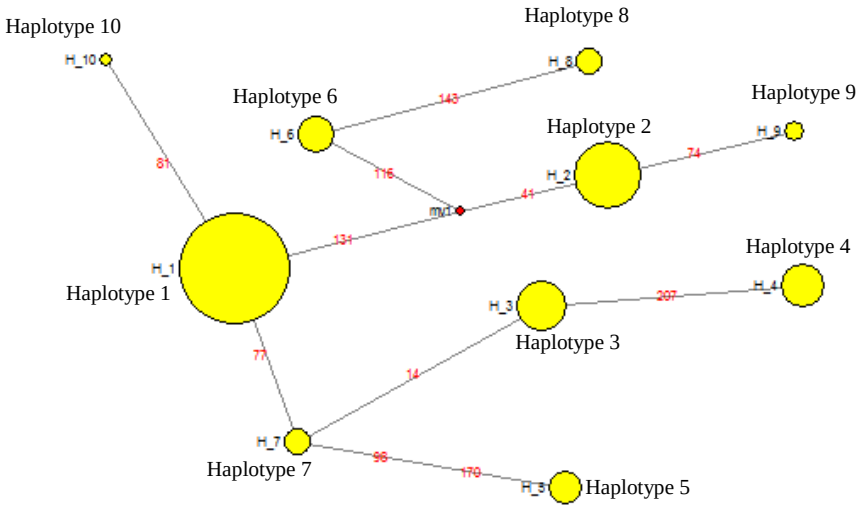


Figure 16 - Median-joining network of *T. sanguisuga* haplotypes. Default assumptions were utilized for network construction and standard post-processing was used to simplify the network. The network diagram displays the same topology as the maximum likelihood tree.

To examine the movements of haplotypes, spatial data were plotted in ArcMap 10.4 for Desktop (ESRI). Haplotype 1 was chosen for hot-spot analysis (Figure 17) because it was the most abundant haplotype identified. The locations of Haplotype 1 bugs were analyzed for hot- and cold-spots with respect to the year of collection. Because the year of collection was used as the stratifying variable, cold-spots represent clustering in less recent years and hot-spots represent clustering in more recent years. The hot-spot analysis suggests that the cytochrome b locus is informative of the movement of genetically related *T. sanguisuga* groups through space at this small scale. Statistically significant spatial clustering Haplotype 1 bugs occurs in early collection years (approximately 2008-2010) and late collection years (approximately 2013-2015) at different ends of the property. Of particular interest is the observation that in 2012, a large construction project began adjacent to the area where Haplotype 1 was found in the early collection years. During the initial land clearing for the project, construction workers stated that they removed a large amount of material that appeared to be rodent nests and observed many rodents during the process. That year, only two specimens of Haplotype 1 were collected, despite no significant change in the total number of bugs collected (of any haplotype) in the preceding or following year. By comparison, eight bugs of Haplotype 1 were collected in 2011 and ten bugs of Haplotype 1 were collected in 2013.

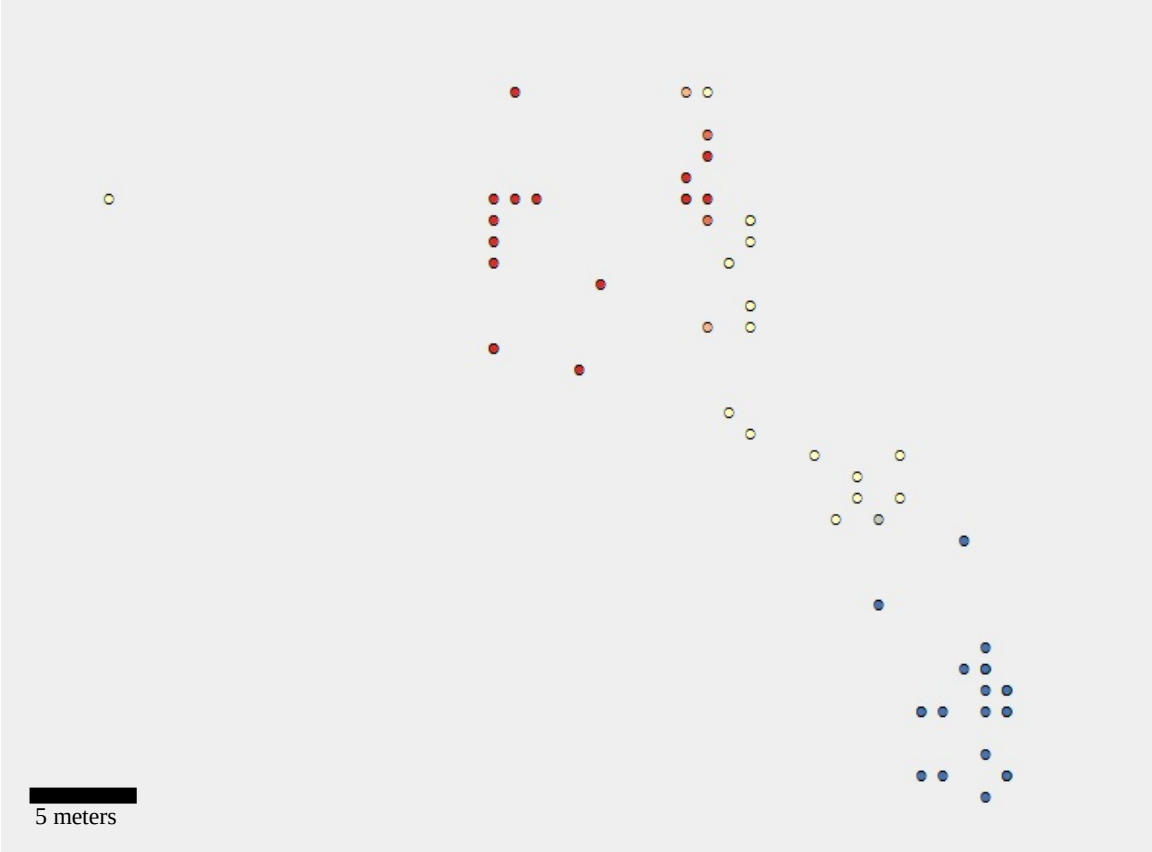


Figure 17 - HotSpot analysis of Haplotype 1 by year of collection. Five levels of significance are reflected by the colored locations. Light blue and dark blue represent 80% and 90% confidence of a cold spot, respectively. Light red and dark red represent 80% and 90% confidence of a hot spot, respectively. Yellow locations were not statistically significant.

This observed change in spatial clustering by haplotype suggests that genetically related *T. sanguisuga* live near one another over time periods covering multiple generations of offspring. This assumes that the relatively short flight range of *T. sanguisuga* [5] results in each bug traveling the shortest distance possible to arrive at the collection location. Spatial clustering by haplotype is important because it suggests that *T. sanguisuga* leads a largely sedentary existence. Such a life would require an animal, as an obligate bloodmeal source, to reside in a single space for multiple years or produce multiple generations in the same space for several years. Additionally, these data suggest that vector control efforts targeting *T. sanguisuga* living around such an animal would be highly effective.

The present research is the first report of small-scale population genetics of *T. sanguisuga*. A significant amount of work demonstrating the movements of sub-populations of *T. dimidiata* in the Yucatan Peninsula [74] and *T. sordida* and *Panstrongylus geniculatus* in Brazil [77] has been reported previously. These reports are concerned with the movement of populations of peri-domestic and sylvatic species around and into human dwellings rather than population structure in the sylvan environment. The small-scale population movements of *Rhodnius* spp. have been investigated in Ecuador and were found to be associated with squirrels and small rodents, and the movements of these bugs have been demonstrated to be the result of the movement of its associated animal [78, 79]. The present research suggests that the movements of *T. sanguisuga* may similarly depend on the movement of an associated animal.

It should be noted that bugs of Haplotype 1 were the most common haplotype collected. Because of this haplotype's proportional abundance, it is possible that *T. sanguisuga* populations move frequently and that clustering was a matter of chance. However, given the detection of ten haplotypes at this single residence, it is unlikely that the genetic locus lacks the resolution to preclude such confounding.

Like the large-scale analysis, the present research is limited by the relatively short size of the amplified gene fragment. Additionally, the low yield of DNA suitable for analysis limited the present research. The poor recovery from collected bugs is likely the result of poor condition at collection (often dead and desiccated for and unknown amount of time) and inconsistent storage practices among laboratory staff. Further field collections of specimens from the vicinity of an associated animal would be helpful in overcoming this limitation.

T. cruzi discrete typing unit (DTU) and spatial correlation

Understanding that haplotypes of *T. sanguisuga* spatially cluster, the question of *T. cruzi* DTU was investigated to determine what, if any, relationship exists between different *T. cruzi* DTUs and haplotype. By monitoring the prevalence of different *T. cruzi* DTUs within different haplotypes of *T. sanguisuga,* inferences may be made regarding the natural sylvatic cycle of the parasite.

Samples that tested positive for *T. cruzi* were subjected to further analysis to determine the DTU of the *T. cruzi* DNA in the sample. Of the 428 samples of DNA from *T. sanguisuga* abdomens from Site 1, 232 were determined to be *T. cruzi* positive by demonstrating amplification with the D71/D72 primer set (Figure 18). These positive samples were further analyzed by HRMA with the V1/V2 primer set (Figure 19) using conservative software constraints to identify similar sequences.

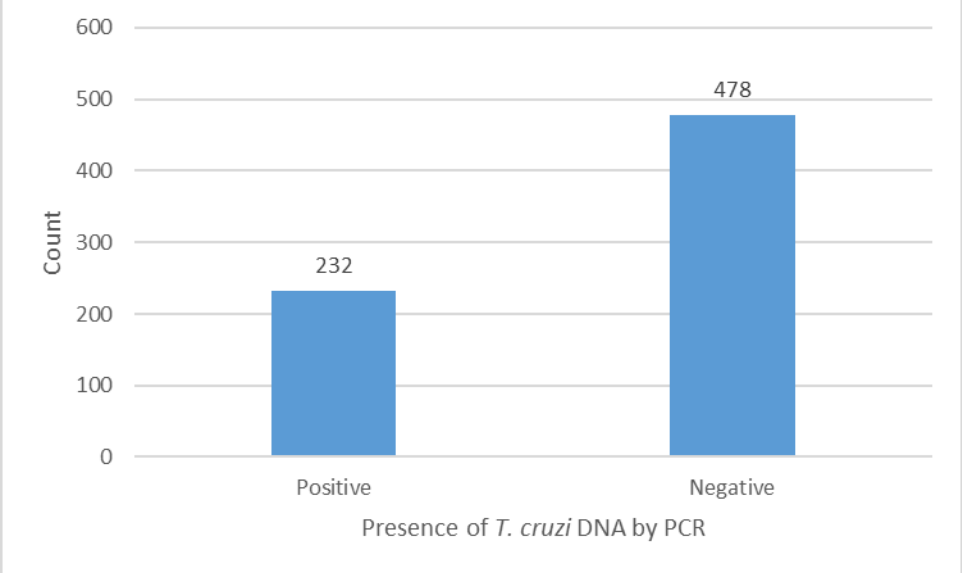


Figure 18 - The D71/D72 primer set was used to screen the DNA samples extracted from *T. sanguisuga* abdomens. The positive samples identified by this screen were further analyzed to determine the discrete typing unit (DTU) of the *T. cruzi* parasites.

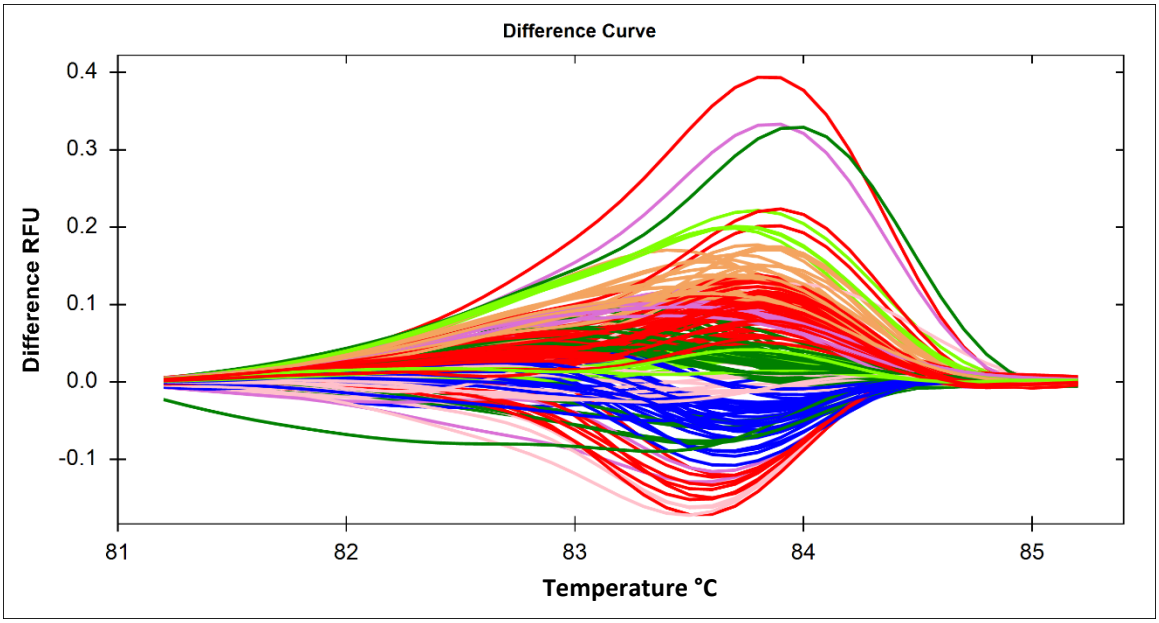


Figure 19 - Samples identified by the D71/D72 screen were analyzed by HRMA with the primer set V1/V2. A representative sample from each software-called sequence was sent for sequencing. Each color represents a different sequence identified by the HRMA software.

Sequences were submitted to the TriTryp database [80] using its PopSet dataset for comparison. All twelve clusters called by the software were identified as *T. cruzi* by the database and matched reference strains of known DTU (Table 10) [81]. Two of the six DTUs of *T. cruzi* were detected. These two DTUs, Tc I and Tc IV, have both been previously described in southeastern Louisiana [8]. In contrast to previously published works about *T. cruzi* in *T. sanguisuga* [8], the predominant DTU identified is Tc IV (59%) (Figure 20). Other studies in the United States have found Tc IV in *T. sanguisuga* and sylvatic animals [82, 83]. Of these animals raccoons have been proposed as a wildlife reservoir of this DTU [83]. Of public health importance are the differing pathologies of these two DTUs. Tc I is associated with cardiomyopathy [84] while Tc IV is associated with acute Chagas disease after oral exposure [85]. Because the majority of the samples were collected at a single residence (Figure 21), it is possible that Tc IV is over-

represented. However, it is interesting that only Tc IV was detected north of the Interstate-10 corridor (Figure 21). Although the absence of Tc I in this area is likely a product of small sample size, the fact that only Tc IV was found in these samples suggests that the observed high relative prevalence of Tc IV is likely representative of the region's parasite population. The DTU of the samples was compared with haplotype data, but there was no relationship found between haplotype and *T. cruzi* DTU. This may suggest that Tc I and Tc IV in southeastern Louisiana do not exist in separate sylvatic cycles or that the cytochrome b haplotype of *T. sanguisuga* is not associated with a specific animal or sylvatic cycle.

Table 10 – Database identification of *T. cruzi* DTUs.

Cluster	Top PopSet Match	DTU Determination
1	CA-1	Tc I
2	91122102	Tc IV
3	91122102	Tc IV
4	91122102	Tc IV
5	CA-1	Tc I
6	91122102	Tc IV
7	91122102	Tc IV
8	Honduras Clone 2	Tc I
9	CA-1	Tc I
10	CA-1	Tc I
11	91122102	Tc IV
12	91122102	Tc IV

Sequences from the 12 clusters identified by the software were sequenced and checked compared to the PopSet dataset in the TriTryp database [80]. All samples returned sequences from *T. cruzi.* The samples matching most closely with the sequence were identified and researched to determine the assigned DTU [81].

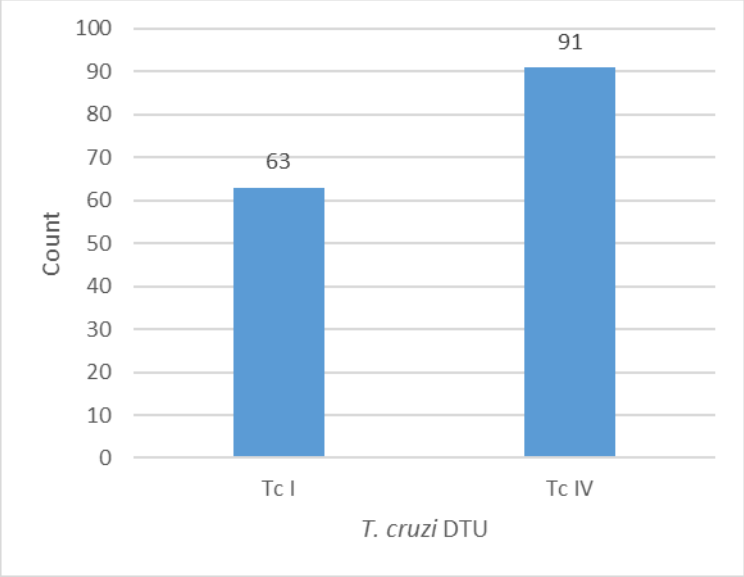


Figure 20 - After classification of each of the identified sequences, it was determined that 59% (n=91) of the samples belonged to the Tc IV DTU and 41% (n=63) belonged to the TcI DTU.

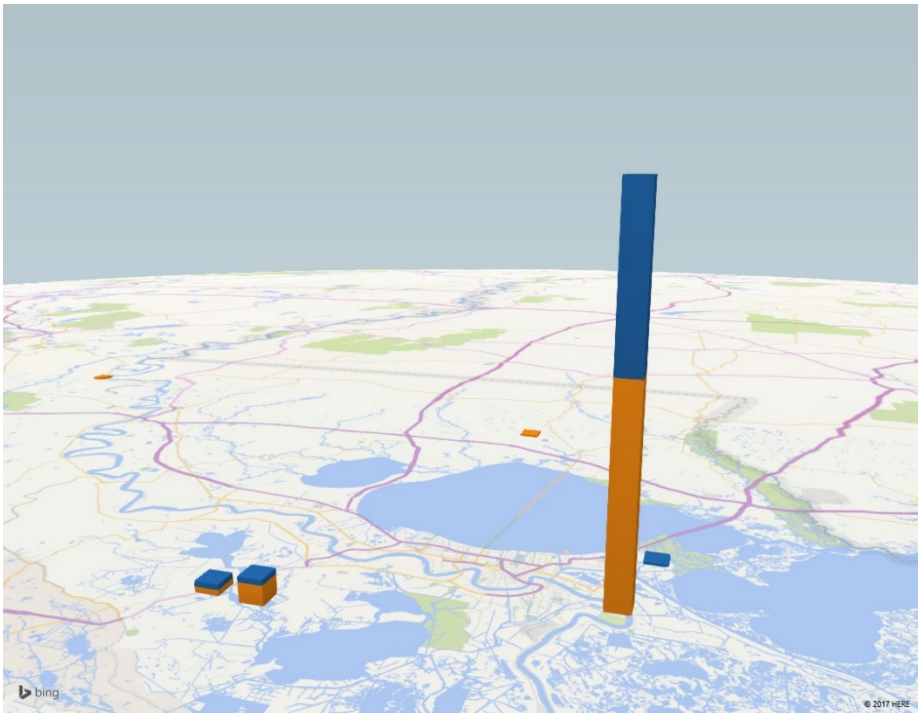


Figure 21 - The distribution of *T. cruzi* DTUs identified in southeastern Louisiana. The size of each graph represents the number of samples at that location. Blue represents Tc I and orange represents Tc IV. The samples collected in Orleans Parish are represented in aggregate for clarity.

Chapter VI

Aim 2 – Results and Discussion

Identification of bloodmeal clusters by HRMA

To gain insight into the vertebrate sources of blood for *T. sanguisuga* in southeastern Louisiana, HRMA was used to identify different sequences of the vertebrate cytochrome b gene within the abdominal contents of collected *T. sanguisuga*. While there were many abdominal DNA samples (n=710), only 314 demonstrated amplification using conventional PCR. When these samples were re-assayed using HRMA, 195 produced satisfactory melt curves (Figure 22). Current HRMA technology is not designed to detect multiple sequences in a single reaction. However, such instances are readily observable in the negative derivative graphs of melt curve data. To identify melt curves demonstrating multiple cytochrome b sequences, negative derivative melt curve data were exported from the CFX Manager software to Microsoft Excel (Figure 23). Curves were matched manually by melting temperature and curve shape.

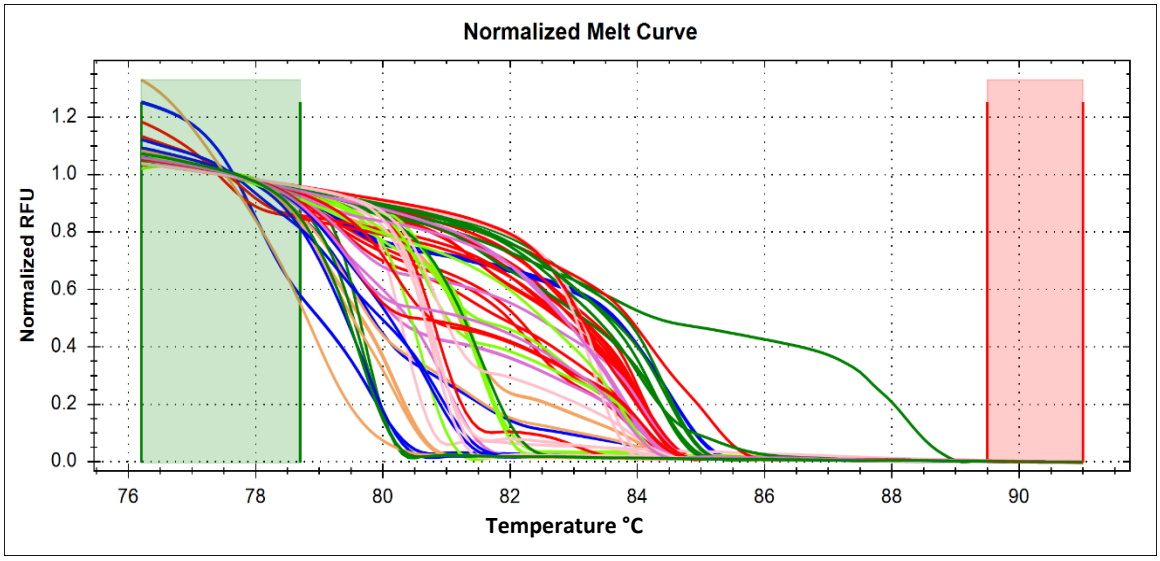


Figure 22 - Normalized HRMA melt curves of the cytochrome b gene from *T. sanguisuga*. Each color represents a unique sequence as determined by the software. Due to the large variability in the sequences, each suspected cluster was analyzed with similar sequences to allow for appropriate adjustment of the normalization ranges for pre- and post-melt data

.

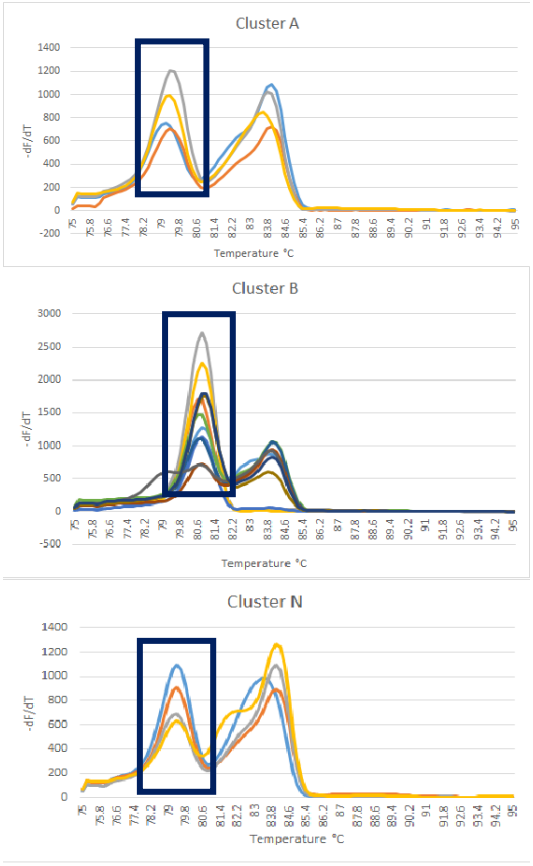


Figure 23 - Negative derivative melt curve data generated by the CFX Manager software was copied into Microsoft Excel and graphed. Melting temperature (indicated by the peak of each curve) and shape of the curve were used to group samples with multiple cytochrome b sequences. Black boxes indicate unique sequences identified and grouped together.

The nucleotide sequence of a representative sample of each cluster was identified by nucleotide BLAST search (NCBI) (Table 11 and Figure 24).

While many vertebrates have been previously identified as bloodmeal sources of *T. sanguisuga* (Table 2), the present research aims to determine the frequency with which *T. sanguisuga* feeds on the eastern woodrat and humans. The research's focus on the eastern woodrat was predicated on field observations of *T. sanguisuga* and the known associated animals of closely related triatomine species. Two specific field observations contributed significantly to the research's focus on the eastern woodrat. As previously mentioned, a construction event at Site 1 disturbed many rodents living in piles of limbs and fallen trees. Based on previous rodent trapping at Site 1 [8], the eastern woodrat was the only rodent collected from the site that naturally inhabits such environments. The other species collected (the cotton mouse, *Peromyscus gossypinnus*, and the house mouse, *Mus musculus*) were only found inhabiting the home. Neither species is known to construct large middens, or homes, from fallen trees. The midden-building habits of the eastern woodrat have been well documented in Louisiana [86] and are consistent with the descriptions provided by the construction workers. The second observation occurred in the Bayou Sauvage National Wildlife Refuge in New Orleans, Louisiana. The outer housing of a dilapidated bee hive was found to contain many *T. sanguisuga* at every life stage. Beneath the raised bottom of this housing was rodent bedding material. The housing was deconstructed during collection of *T. sanguisuga* specimens. The site was monitored weekly for one month. During that time, there appeared no evidence of a rodent cleaning or attempting to re-inhabit the space. However, the aggregation of *T. sanguisuga* remained in the same location with no observed dispersal during the

observation window. These observations are consistent with observations of other triatomine species in the United States which often find multiple life stages of these species inhabiting the middens of woodrat species [46, 87].

In contrast to the field observations, none of the collected *T. sanguisuga* contained blood from an eastern woodrat. Only one sample of the 195 identifiable vertebrate cytochrome b sequences contained rodent DNA, and this DNA belonged to a marsh rice rat (Table 11). The most often identified vertebrate DNA sequence in the samples was human. Non-human sources of vertebrate blood were pig (n=11), cow (n=9), domestic dog (n=5), raccoon (n=4), and domestic cat (n=1) (Figure 24).

T. sanguisuga collected alive from the field were often kept alive in the laboratory to establish a laboratory colony. During this process, these bugs were fed bloodmeals through a membrane feeding device. Over the years human, pig, and cow blood were all used for this purpose. No record was kept of which bugs were fed in this way; it is not possible to determine which of these samples may have been the result of natural feeding and which may have been the result of laboratory feedings. It should be noted, however, that two of the samples containing cow DNA were collected from a cow pasture and that wild pigs are endemic to the region.

The detection of dogs as a bloodmeal source is important, but not unexpected. There have been many reports of canine Chagas disease across the southern United States [25, 44, 88, 89]. The most probable route of *T. cruzi* infection in dogs is ingestion of triatomine bugs [25]. Because canine Chagas disease often manifests as myocarditis with conduction abnormalities [44], canine Chagas disease may be contributed to heartworm and, thus, be underreported. This is important to human health as previous research has

demonstrated that dogs exhibit a long-lived parasitemia that makes them efficient peridomestic reservoir hosts for the parasite [90–92]. Likewise, the detection of raccoons as a bloodmeal source is in agreement with a previously published account [32], and many studies have found wild raccoons in the southern United States to be naturally infected with *T. cruzi* [45, 83, 93]. The presence of domestic cat blood is interesting in that relatively few researchers have studied the importance of cats, as opposed to dogs, as a peri-domestic reservoir of *T. cruzi* [13, 90]. The present research is the first description of the marsh rice rat as a bloodmeal source for any triatomine species. This specimen was collected at Site 1. The marsh rice rat was not detected during a peri-domestic rodent trapping study at this location [8], however this species is not known to thrive in a peri-domestic habitat and has been collected by others in the area [86].

Table 11 - Vertebrate sequences identified by BLAST search.

HRMA Cluster	Species	Common Name
Cluster A	*Bos taurus*	Cow
Cluster B	*Sus scrofa*	Pig
Cluster C	*Sus scrofa*	Pig
Cluster D	*Homo sapiens*	Human
Cluster E	*Sus scrofa*	Feral hog
Cluster F	*Procyon lotor*	Raccoon
Cluster G	*Felis catus*	Domestic cat
Cluster H	*Oryzomys palustris*	Marsh Rice Rat
Cluster I	*Procyon lotor*	Raccoon
Cluster J	*Canis lupus familiaris*	Dog
Cluster K	*Canis lupus familiaris*	Dog
Cluster L	*Sus scrofa*	Pig
Cluster M	*Canis lupus familiaris*	Dog
Cluster N	*Bos Taurus*	Cow
Cluster O	*Homo sapiens*	Human
Cluster P	*Bos Taurus*	Cow
Cluster Q	*Homo sapiens*	Human
Cluster Z	*Homo sapiens*	Human

Clusters that were found to contain the same species as another cluster were found to be samples containing more than one cytochrome b sequence or over-classified by the conservative settings of the software.

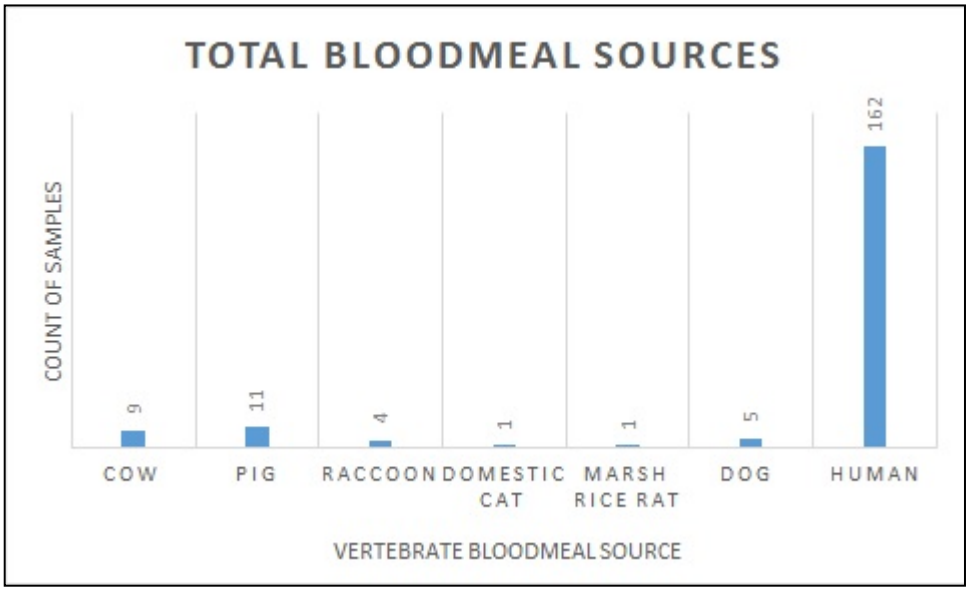


Figure 24 - Number of samples amplifying vertebrate sources of cytochrome b by species common name.

Identification of probable DNA contamination

Human DNA was detected in many of the samples (n=162). Although previous bloodmeal analyses of *T. sanguisuga* have found human DNA and suggested it is the direct result of feeding [29, 32], the present research sought to determine whether or not human DNA in these samples is the result of contamination. To assess this possibility, the generated cytochrome b melt curves from the original HRMA assay (Figure 25) were re-analyzed using the most conservative software settings – resulting in 7 distinct software-identified groups (Figure 26). Ct values for the human samples were not appreciably different than those from non-human animals and positive controls. The genomic DNA of an individual heavily involved in the research was extracted and analyzed with these newly identified groups. This individual's DNA was determined by the software to match the DNA from four of the *T. sanguisuga* samples. This suggests that human DNA contamination during nucleic acid extraction and processing does occur. The small number (n=7) of sequences detected across multiple collection years (Figure 27) and from multiple collection sites (Figure 28) also suggest a limited number of laboratory contamination events. .

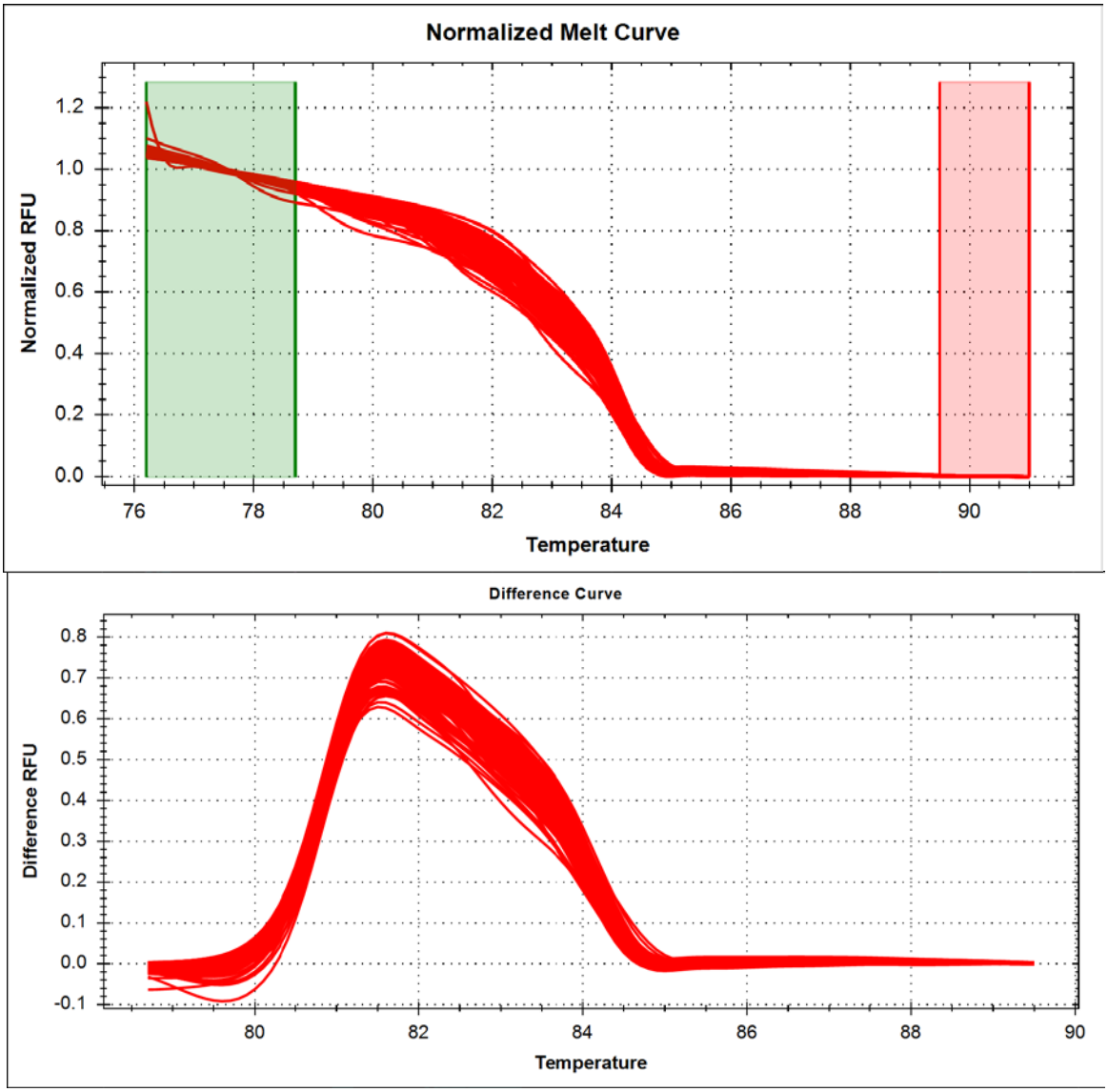


Figure 25 - These are graphical representations of the human cytochrome b sequences as called by the initial experiment. Above: Normalized melt curves. Below: Difference curves of the melt data. All samples determined by the software to be alike (single color).

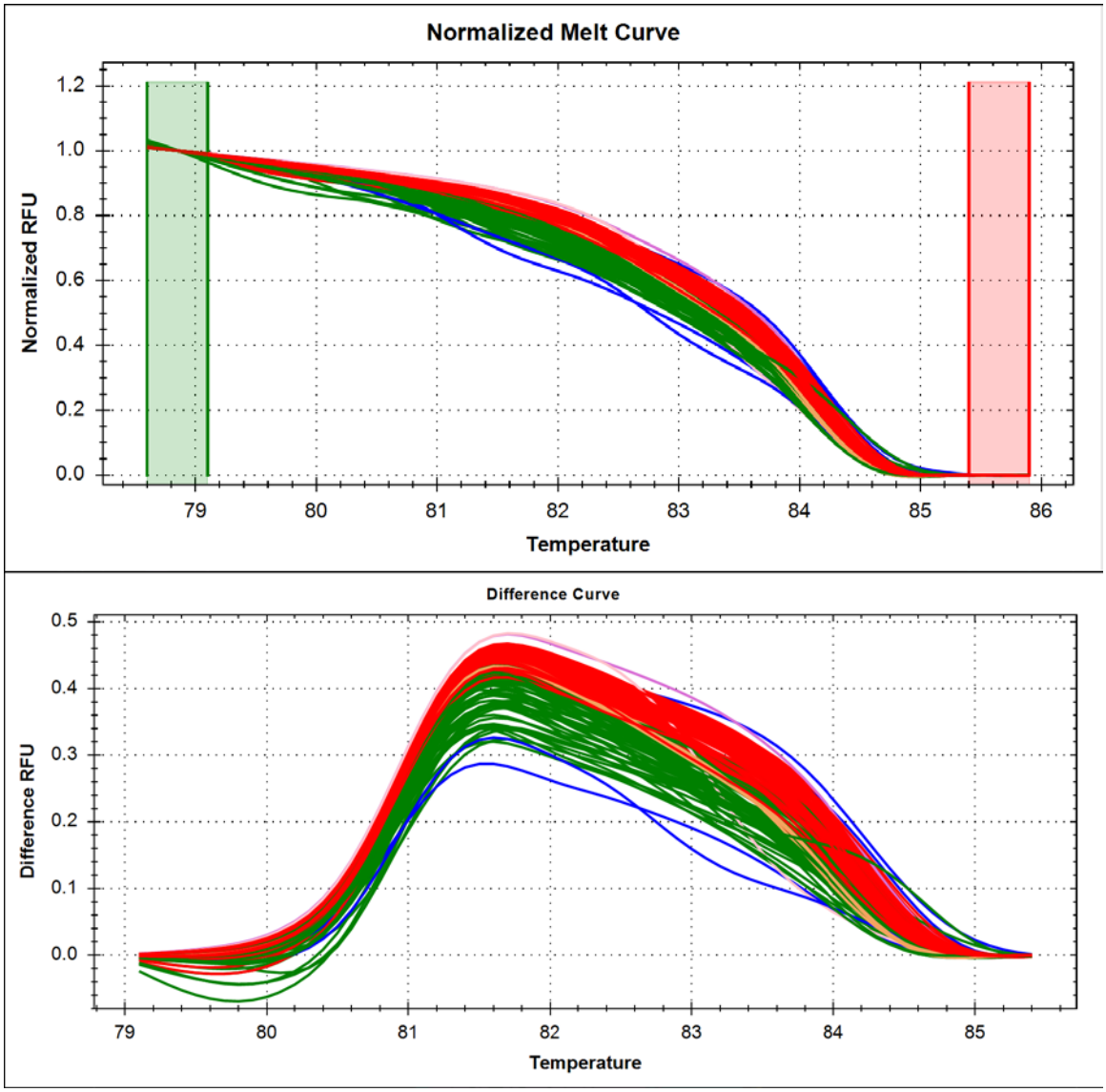


Figure 26 - By increasing the sensitivity settings in the software, 7 distinct human cytochrome b clusters were observed.

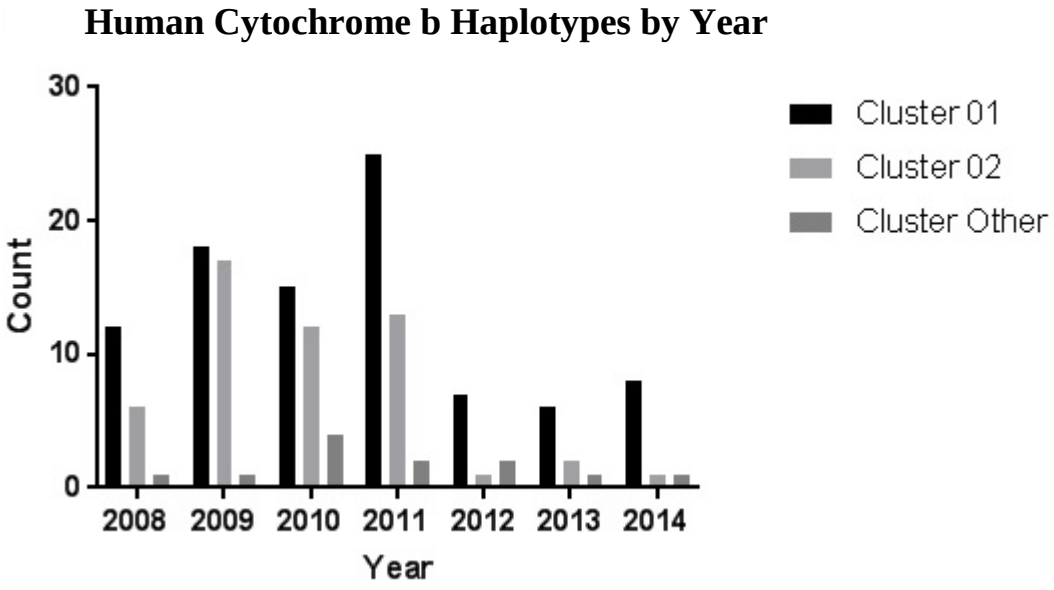


Figure 27 - The human cytochrome b clusters determined by the software were graphed by year to determine any trends.

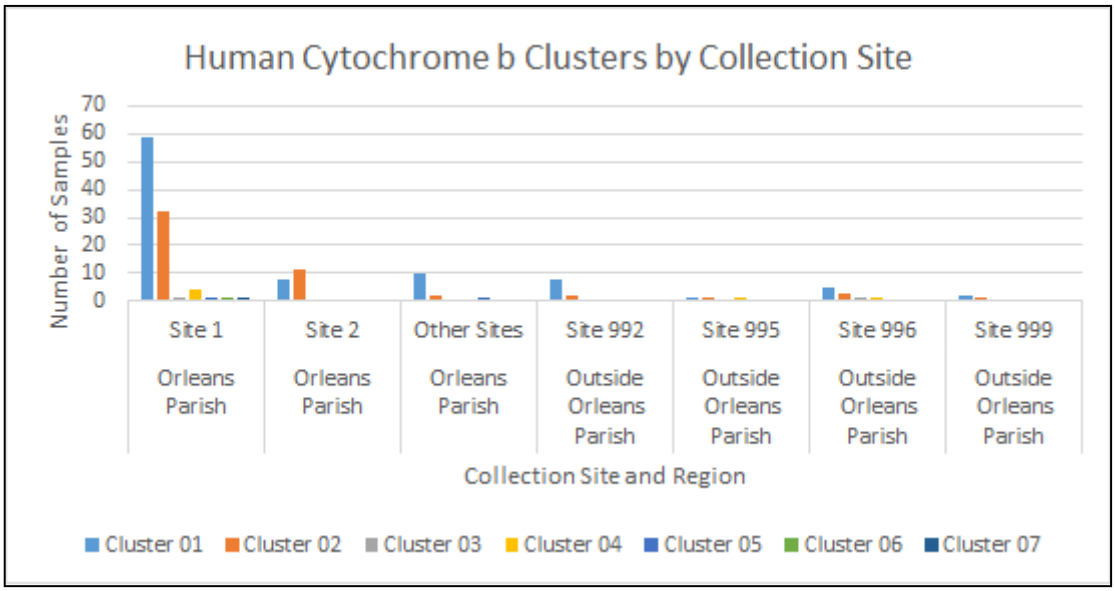


Figure 28 - The human cytochrome b clusters determined by the software were stratified by collection site and region (either in or outside of Orleans Parish).

A major limiting factor was the inability to sequence the human DNA samples due to privacy concerns. Without these data, it is impossible to determine if the amplified locus is variable enough to separate non-related individuals from one another. Previous studies have determined that the human cytochrome b gene is sufficient, in most cases, to identify non-related individuals [94, 95]. However, these studies identified haplotypes using the entire cytochrome b gene (1.1kb). This limitation could be overcome by high resolution electrophoresis of the amplified and enzyme-digested D-loop regions of the mitochondrial DNA as others have [96]. Unfortunately, the equipment and resources were not available to use this technology in the present research

Chapter VII

Aim 3 – Results and Discussion

Maximum entropy modeling of *T. sanguisuga*

The development of a MaxEnt model for the suitable habitat of *T. sanguisuga* was necessary as the most recent distribution model for this species [4] was constructed using a small number of unique data points (n=18) and did not predict the presence of *T. sanguisuga* throughout much of southeastern Louisiana. Publicly available *T. sanguisuga* presence data from iDigBio [62], iNaturalist [63], and the Louisiana State Arthropod Museum [97] were used to generate the model. iDigBio aggregates and maintains records from contributing institutions. iNaturalist is an online community where citizen scientists can report presence data of plants and animals. For this study, only samples with pictures that could be identified as *T. sanguisuga* were used. The final model (Figure 29) closely resembles the distribution map published by the Centers for Disease Control and Prevention (Figure 1) including a predicted absence in West Virginia. Using multiple environmental data layers commonly used for this purpose [4, 70, 79, 98] (Table 7), the MaxEnt algorithm determined the presence of *T. sanguisuga* to be positively associated with elevated temperatures, high rainfall, and moderate contiguous leaf coverage (Table 12). One interesting aspect highlighted by the model is the association of T. sanguisug*a* presence with precipitation levels. This is interesting because *T. sanguisuga* in the field are not found in wet or damp environments [23]. It is likely, then, that precipitation and *T. sanguisuga* presence are correlated with a currently unidentified third variable that is tightly associated with precipitation. These findings are significant because they demonstrate that the MaxEnt program creates models that predict suitable habitat for *T.*

sanguisuga across southeastern Louisiana where previous models have predicted unsuitable habitat.

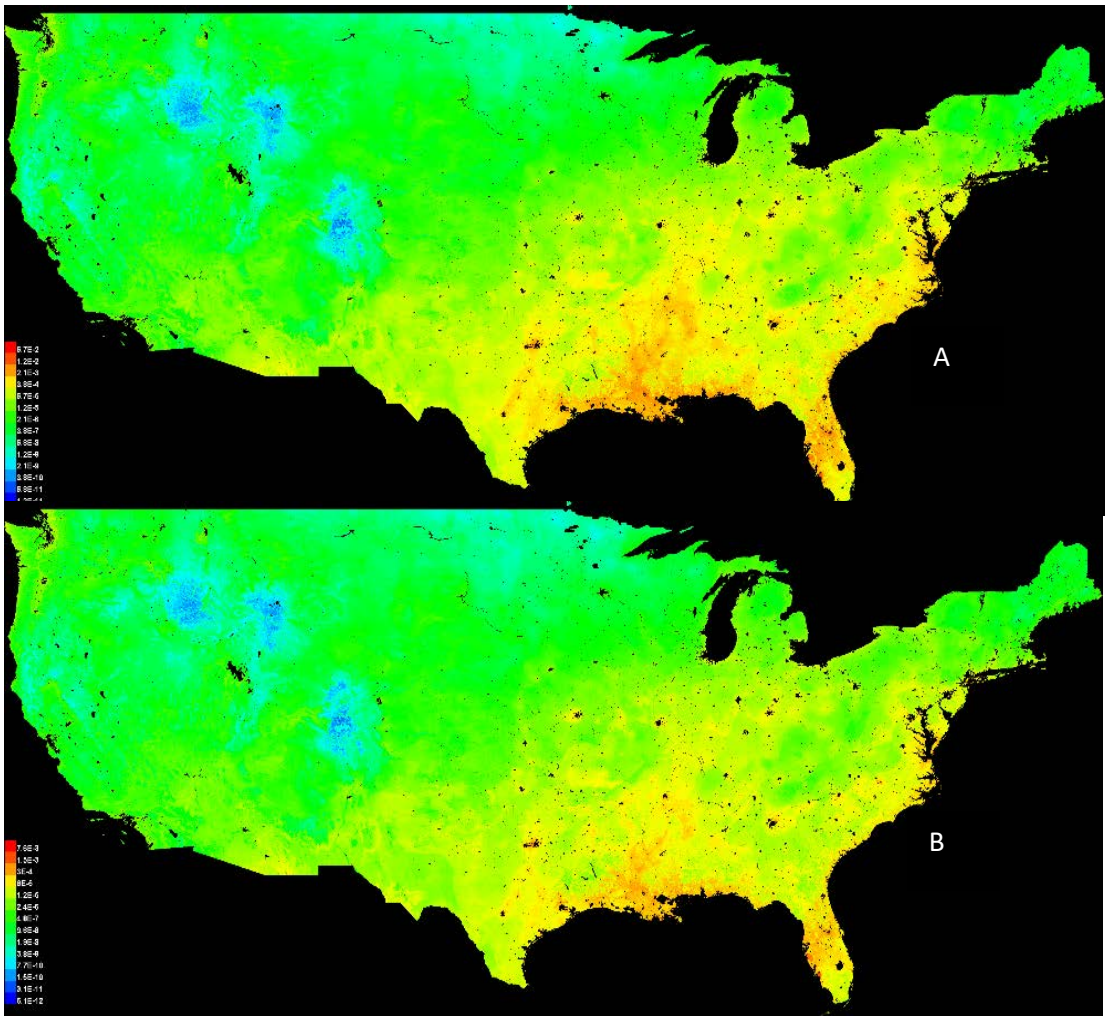


Figure 29 - MaxEnt model generated for *T. sanguisuga* using 50 cross-validated iterations. (A) Mean prediction of suitable habitat. (B) Standard deviation of the cumulative model iterations.

Table 12 - Jackknife results of the variables contributing to the habitat model of *T. sanguisuga.*

Rank	Variable	Percent contribution	Permutation importance
1	Mean temperature of warmest quarter	18.2	10.9
2	Precipitation in driest month	13.9	22.8
3	Annual mean temperature	11.2	0.8
4	Mean diurnal temperature range*	(9.7)	10.3
5	Precipitation of driest quarter	6.7	1.6
6	Precipitation of wettest quarter	6.5	0.1
7	Average maximum leaf area index	(5.4)	12
8	Minimum temperature of coldest month	5.2	13.7
9	Precipitation of warmest quarter	5.2	4.2
10	Average minimum leaf area index	4.5	3.2
11	Annual temperature range	(3.5)	1.6
12	Mean temperature of driest quarter	1.7	2.2
13	Annual precipitation	1.4	0.5
14	Maximum temperature of wettest month	(1.1)	0.2
15	Precipitation of coldest quarter	1	1.5
16	Mean temperature of wettest quarter	0.9	0.2
17	Landcover classification	0.9†	0.6
18	Mean temperature of coldest quarter	0.8	0.9
19	Precipitation of wettest month	(0.7)	4.7
20	Isothermality**	(0.6)	0.3
21	Precipitation seasonality***	0.4	7.2
22	Temperature seasonality****	(0.3)	0.5

*The mean diurnal temperature range is calculated as the mean of monthly maximum and minimum temperatures. ** Isothermality is calculated as the mean diurnal temperature range divided by annual temperature range. ***Precipitation seasonality is the coefficient of variation of precipitation among the seasons. ****Temperature seasonality is the standard deviation of temperature by season. † The most positively associated landcover classifications were mixed forests, woodlands, and deciduous broadleaf forests. Contribution percentages enclosed in parentheses indicate a negative association.

Distributions of associated animals

MaxEnt modeling of suitable habitats for *T. sanguisuga* was undertaken to better understand the ecological niche of this vector and the overlap of its suitable habitat with that of the eastern woodrat. The present research focuses on the eastern woodrat because of the previously described field observations.

MaxEnt models were created for animals reported to be associated with *T. sanguisuga* (Table 8). Additionally, a MaxEnt model was created for *Oryzomys palustris* and included in the final analysis as this species was identified as a bloodmeal source in the course of research (Figure 30). Presence data for *O. palustris* was obtained as previously described from the Integrated Digitized Biocollection Database [62] (Figure 29).

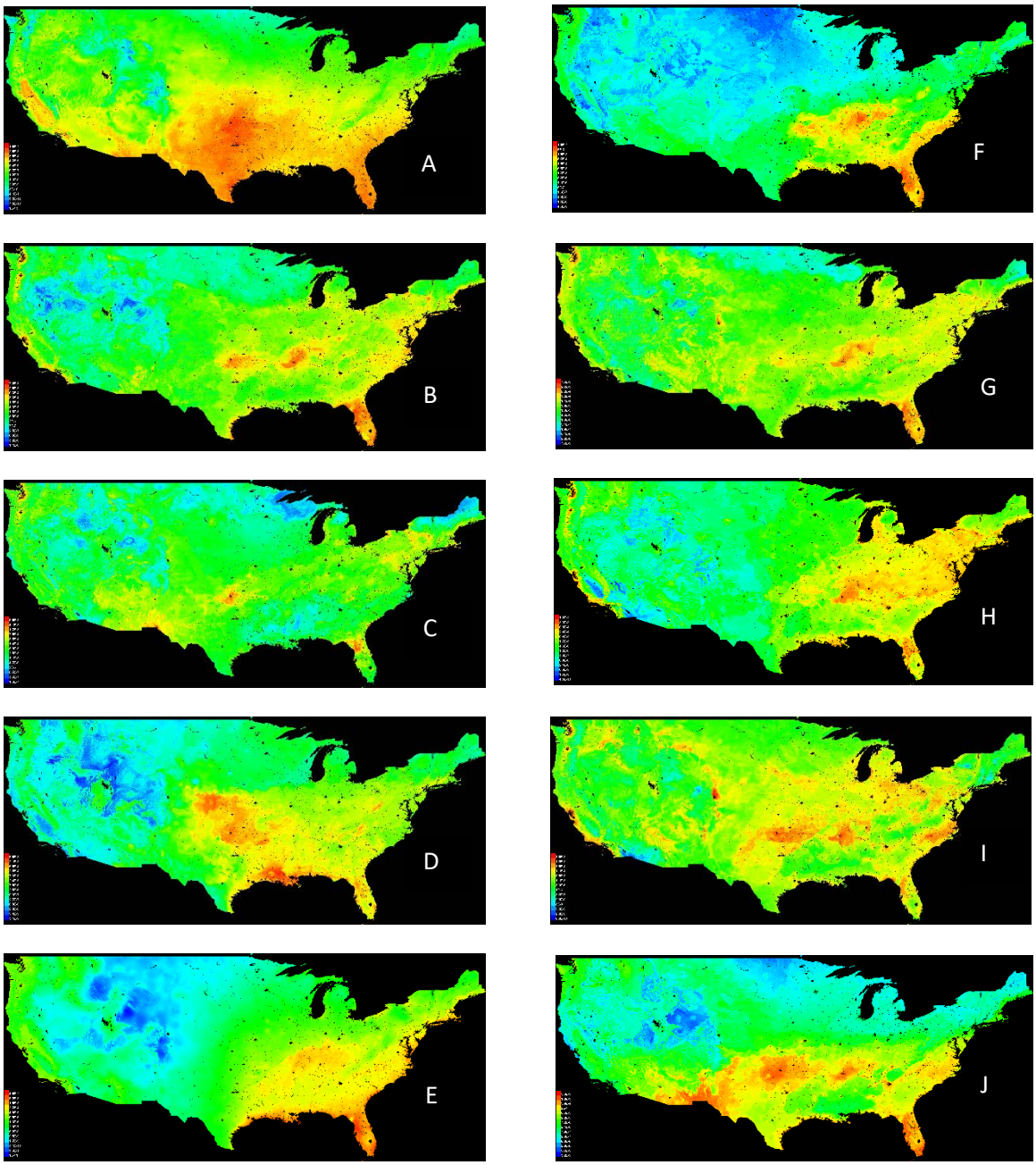


Figure 30 - Average MaxEnt models for animals associated with *T. sanguisuga*. (A) Nine-banded armadillo (B) Virginia opossum (C) Striped skunk (D) Eastern woodrat (E) Marsh rice rat (F) Cotton mouse (G) Raccoon (H) Eastern gray squirrel (I) Fox squirrel (J) Hispid cotton rat.

Modeling *T. sanguisuga* from associated animals' distributions

In order for any animal to be closely associated with an insect with a low dispersal range [5], the two species must be able to inhabit the same geographic areas. Based on this premise, MaxEnt models of animals associated with *T. sanguisuga* were used as possible predictive variables for habitat suitability for *T. sanguisuga*. The jackknife calculation for ranking variable importance was used to determine the relative importance of each species' suitable habitat or distribution with respect to the suitable habitat modeled for *T. sanguisuga*. By comparing MaxEnt models of suitable habitat for *T. sanguisuga* with those of reportedly associated animals, the present research creates a framework by which potentially associated animals of *T. sanguisuga* can be ranked from the most to least likely associated animal. Utilizing this framework, the marsh rice rat and eastern woodrat were consistently ranked as the two most closely associated animals of *T. sanguisuga* out of the animals examined (Table 13).

Table 13 - Contribution of each animal's suitable habitat to a MaxEnt model of *T. sanguisuga.*

Rank	Variable	Percent contribution	Permutation importance
1	Marsh rice rat	37.5	16.2
2	Eastern woodrat	21.3	6.7
3	Nine-banded armadillo	20.7	34.2
4	Cotton mouse	8.4	11.3
5	Striped skunk	2.9	3.7
6	Fox squirrel	2.6	4.5
7	Racoon	2.2	5
8	Eastern grey squirrel	1.6	5.4
9	Virginia opossum	1.5	4.4
10	Hispid cotton rat	1.4	8.6

All associations were positive.

As a second means of habitat suitability comparison the final MaxEnt model for *T. sanguisuga* was compared pair-wise with each MaxEnt model for the potentially associated animals using ENMtools software. The species' models that were found to be most overlapping, or most similar, to the suitable habitat of *T. sanguisuga* were the marsh rice rat and the eastern woodrat (Table 14).

Table 14 - Pair-wise statistics calculated by ENMtools.

Species	**Schoener's D**	***I* Statistic**	**Association Rank**
Marsh rice rat	0.462832327	0.748665196	1
Eastern woodrat	0.456507736	0.737523789	2
Virginia opossum	0.444760717	0.726123938	3
Cotton mouse	0.393052109	0.676295184	4
Raccoon	0.383242718	0.664092918	5
Hispid cotton rat	0.379615029	0.654069263	6
Eastern gray squirrel	0.377142345	0.667722562	7
Nine-banded armadillo	0.342882686	0.658527764	8
Fox squirrel	0.333967927	0.615222159	9
Striped skunk	0.252473478	0.530421253	10

Calculation of model overlap between MaxEnt models of the indicated species and the mean MaxEnt distribution of *T. sanguisuga*. Each statistic spans from 0 (no overlap between the two models) to 1 (complete agreement of the two models). Values have been ranked based on their *D* and *I* statistics.

Although both methods of distribution comparison identified the same two rodents as the species with the most similar habitat to that of *T. sanguisuga*, it is interesting that the methods of comparison begin to disagree greatly as the compared distributions become more dissimilar. The primary tool used in the literature for comparing distribution and habitat models of triatomine vectors is ENMtools [70]. The use of MaxEnt generated models as predictive layers within MaxEnt has not been established in the literature in the same way. However, the agreement between the comparison tools with regard to the most similar distributions is interesting and suggests that this method may be a useful in future studies.

The MaxEnt models in the present research are limited by available collection data. Some areas of the models may appear to be more significant due to oversampling. These oversampled areas tend to be located near urban centers. Indeed, initial MaxEnt models found *T. sanguisuga* so highly correlated with urban and built environments that the importance of other variables was difficult to discern. To control for this, areas of urban and built environment were removed a priori from all layers prior to running the final models. From what is known of about the biology of *T. sanguisuga*, it is unable to live in urban and built environments. Although adults are periodically found inside of homes, there are no reports that this species successfully colonizing a home built with modern construction methods. The close association observed between urban centers and *T. sanguisuga* is presumed to be the result bias in sampling effort created by research universities and passive surveillance for these bugs by suburban homeowners.

Chapter VIII

Conclusions

The present research supports the current theory of two populations of *T. sanguisuga* in southeastern Louisiana. The cytochrome b gene has been demonstrated to be useful in demonstrating the movement of a genetically related group of these bugs at the home of the sixth case of autochthonous human Chagas disease in the United States. This genetic locus has been shown to not be informative of the genetics of the population at a larger scale, though. To characterize the movement of populations across large distances, new genetic markers will need to be characterized. This process will undoubtedly be complicated by the lack of a reference genome. Because this work cannot refute the possibility of a subspecies of *T. sanguisuga* in southeastern Louisiana, it is necessary that this avenue continue to be investigated before further studies of population genetics are undertaken – as there may, in fact, be two independent populations. The implications of defining a subspecies in the region would be far-reaching. Field and laboratory experiments would be needed to determine each group's ecological niche and vectoral capacity.

The bloodmeal analysis studies presented here support previous reports of *T. sanguisuga* feeding on raccoons and dogs. Because of the possibility of laboratory contamination of these samples, it remains unclear how frequently, if at all, these vectors fed on humans, pigs, and cows. The identification of the marsh rice rat as a bloodmeal source is a unique finding that is further supported by the MaxEnt models of suitable habitat for the species. The larger question raised by these results is the whether or not bloodmeal analysis is an appropriate method for answering fundamental questions about

the natural history of *T. sanguisuga*. Because the present research and every other bloodmeal analysis study involving this species relies on dispersed adult specimens, it is unlikely to yield meaningful information about the primary associated animal of the vector. While these data, including incidental feeding on humans, are important from a public health perspective, they are not helpful in discerning where these bugs spend the first year of their life. Bloodmeal analysis data are also problematic when reported without an attempt to control for contamination with human DNA. The cumulative effect of such uncontrolled data in the literature is to overstate the frequency of human-bug interactions – creating a public health concern where there is none. By illuminating the full natural history of this vector species, the true frequency and risk of human-bug interactions may be assessed.

The MaxEnt model of suitable habitat for *T. sanguisuga* presented here is the most robust and detailed model to-date for this species. Further refinements to this model may be made by systematic sampling in areas of predicted suitable and unsuitable habitats. This is the first report to predict the location of a triatomine vector based on the suitable habitats of possibly associated animals. The outlined model comparison framework will be used to inform future studies directly targeting the marsh rice rat and eastern woodrat. By ranking the likely association of animals with *T. sanguisuga*, sampling efforts can be narrowly focused where they are the most likely to provide results.

Much about the natural history and population genetics of *T. sanguisuga* in southeastern Louisiana remains unknown. However, the present research has provided

foundational data on the population genetics, blood feeding habits, and habitat of this vector species.

References

1. World Health Organization (2002) Control of Chagas disease. World Health Organ Tech Rep Ser 905:i–vi, 1-109, back cover. doi: 10.1016/S0065-308X(05)61004-4
2. Coura JR, Viñas PA (2010) Chagas disease: a new worldwide challenge. Nature 465:S6–S7.
3. Cesa, K., Caillouet, K.A., Dorn, P.L., Wesson DM (2013) High *Trypanosoma cruzi* (Kinetoplastida: Trypanosomatidae) Prevalence in *Triatoma sanguisuga* (Hemiptera: Redviidae) in Southeastern Louisiana. J Med Entomol 48:1091–1094.
4. Ibarra-Cerdeña CN, Sánchez-Cordero V, Townsend Peterson a., Ramsey JM (2009) Ecology of North American Triatominae. Acta Trop 110:178–186. doi: 10.1016/j.actatropica.2008.11.012
5. Kjos SA, Marcet PL, Yabsley MJ, et al (2013) Identification of bloodmeal sources and *Trypanosoma cruzi* infection in triatomine bugs (Hemiptera : Reduviidae) from residential settings in Texas, the United States. J Med Entomol 50:1126–1139.
6. Olsen PF, Shoemaker JP, Turner HF, Hays KL (1964) Incidence of *Trypanosoma cruzi* (Chagas) in wild vectors and reservoirs in east-central Alabama. J Parasitol 50:559–603.
7. Cantey PT, Stramer SL, Townsend RL, et al (2012) The United States *Trypanosoma cruzi* infection study: Evidence for vector-borne transmission of the parasite that causes Chagas disease among United States blood donors. Transfusion 52:1922–30. doi: 10.1111/j.1537-2995.2012.03581.x
8. Herrera CP, Licon MH, Nation CS, et al (2015) Genotype diversity of

Trypanosoma cruzi in small rodents and *Triatoma sanguisuga* from a rural area in New Orleans, Louisiana. Parasit Vectors 8:123. doi: 10.1186/s13071-015-0730-8

9. Lent H, Wygodzinsky P (1979) Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas disease. Bull Am museum Nat Hist 163:123–520.
10. Galvão C, Carcavallo RU, Da Silva Rocha D, Jurberg J (2003) A checklist of the current valid species of the subfamily Triatominae Jeannel, 1919 (Hemiptera, Reduviidae) and their geographical distribution, with nomenclatural and taxonomic notes. Zootaxa 36:1–36. doi: 10.11646/zootaxa.202.1.1
11. Wesson DM (2017) Unpublished data.
12. Pacheco-Tucuch FS, Ramirez-Sierra MJ, Gourbière S, Dumonteil E (2012) Public street lights increase house infestation by the Chagas disease vector *Triatoma dimidiata*. PLoS One 7:e36207. doi: 10.1371/journal.pone.0036207
13. Moudy RRM, Michaels S, Jameson SB, et al (2014) Factors associated with peridomestic *Triatoma sanguisuga* (Hemiptera: Reduviidae) presence in southeastern Louisiana. J Med Entomol 51:1043–1050. doi: 10.1603/ME13234
14. Kjos S a, Snowden KF, Olson JK (2009) Biogeography and *Trypanosoma cruzi* infection prevalence of Chagas disease vectors in Texas, USA. Vector Borne Zoonotic Dis 9:41–50. doi: 10.1089/vbz.2008.0026
15. Garcia MN, Hotez PJ, Murray KO (2014) Potential novel risk factors for autochthonous and sylvatic transmission of human Chagas disease in the United States. Parasit Vectors 7:311. doi: 10.1186/1756-3305-7-311
16. (2016) Triatomine Bug FAQs. In: Cent. Dis. Control Prev. - Div. Parasit. Dis.

https://www.cdc.gov/parasites/chagas/gen_info/vectors/index.html. Accessed 3 Apr 2017

17. Zingales B, Andrade SG, Briones MRS, et al (2009) A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: Second revision meeting recommends TcI to TcVI. Mem Inst Oswaldo Cruz. doi: 10.1590/S0074-02762009000700021

18. Zingales B, Miles M a., Campbell DA, et al (2012) The revised *Trypanosoma cruzi* subspecific nomenclature: Rationale, epidemiological relevance and research applications. Infect Genet Evol 12:240–253. doi: 10.1016/j.meegid.2011.12.009

19. Ramos-Ligonio A, Torres-Montero J, López-Monteon A, Dumonteil E (2012) Extensive diversity of *Trypanosoma cruzi* discrete typing units circulating in *Triatoma dimidiata* from central Veracruz, Mexico. Infect Genet Evol 12:1341–1343. doi: 10.1016/j.meegid.2012.04.024

20. Sullivan TD, McGregor T, Eads RB, Davis DJ (1949) Incidence of *Trypanosoma cruzi*, Chagas, in *Triatoma* (Hemiptera, Reduviidae) in Texas. Am. J. Trop. Med. 29:

21. Davis DJ, McGregor T, Deshazo T (1943) *Triatoma sanguisuga* (LeConte) and *Triatoma ambigua* Neiva as natural carriers of *Trypanosoma cruzi* in Texas. [U.S. Dept. of Health, Education, and Welfare, Public Health Service, Health Resources Administration]

22. Packchanian A (1940) Natural Infection of *Triatoma heidemanni* with *Trypanosoma cruzi* in Texas. Public Health Rep 55:1300. doi: 10.2307/4583370

23. Grundemann AW (1947) Studies on the biology of *Triatoma sanguisuga* (Leconte)

in Kansas, (Reduviidae, Hemiptera). J Kansas Entomol Soc 20:77–85.

24. Usinger R (1944) The Triatominae of North and Central America and the West Indies and their public health significance. Public Health Bulletin (Wash. D. C). 228:

25. Bern C, Kjos S, Yabsley MJ, Montgomery SP (2011) *Trypanosoma cruzi* and Chagas' disease in the United States. Clin Microbiol Rev 24:655–681. doi: 10.1128/CMR.00005-11

26. de la Rua N, Stevens L, Dorn PLPL, et al (2011) High genetic diversity in a single population of *Triatoma sanguisuga* (LeConte, 1855) inferred from two mitochondrial markers: Cytochrome b and 16S ribosomal DNA. Infect Genet Evol 11:671–677. doi: 10.1016/j.meegid.2014.03.019

27. Monteiro F a., Barrett T V., Fitzpatrick S, et al (2003) Molecular phylogeography of the Amazonian Chagas disease vectors *Rhodnius prolixus* and *R. robustus*. Mol Ecol 12:997–1006. doi: 10.1046/j.1365-294X.2003.01802.x

28. Yong KE, Mubayi A, Kribs CM (2015) Agent-based mathematical modeling as a tool for estimating *Trypanosoma cruzi* vector–host contact rates. Acta Trop 151:21–31. doi: 10.1016/j.actatropica.2015.06.025

29. Gorchakov R, Trosclair LP, Wozniak EJ, et al (2016) *Trypanosoma cruzi* infection prevalence and bloodmeal analysis in Triatomine vectors of Chagas disease from rural peridomestic locations in Texas, 2013-2014. J Med Entomol 53:911–918. doi: 10.1093/jme/tjw040

30. Ibáñez-Cervantes G, Martínez-Ibarra A, Nogueda-Torres B, et al (2013) Identification by qPCR of *Trypanosoma cruzi* lineage and determination of blood

meal sources in triatomine gut samples in México. Parasitol Int 62:36–43. doi: 10.1016/j.parint.2012.09.003

31. Buitrago R, Bosseno M-F, Depickère S, et al (2016) Blood meal sources of wild and domestic *Triatoma infestans* (Hemiptera: Reduviidae) in Bolivia: Connectivity between cycles of transmission of *Trypanosoma cruzi*. Parasit Vectors 9:214. doi: 10.1186/s13071-016-1499-0

32. Waleckx E, Suarez J, Richards B, Dorn PL (2014) *Triatoma sanguisuga* blood meals and potential for Chagas disease, Louisiana, USA. Emerg Infect Dis 20:2141–2143. doi: 10.3201/eid2012.131576

33. Pung OJ, Spratt J, Clark CG, et al (1998) *Trypanosoma cruzi* infection of free-ranging lion-tailed macaques (*Macaca silenus*) and ring-tailed lemurs (*Lemur catta*) on St. Catherine's Island, Georgia, USA. J Zoo Wildl Med 29:25–30.

34. Buitrago NLR, Bosseno MF, Waleckx E, et al (2013) Risk of transmission of *Trypanosoma cruzi* by wild *Triatoma infestans* (Hemiptera: Reduviidae) in Bolivia supported by the detection of human blood meals. Infect Genet Evol 19:141–144. doi: 10.1016/j.meegid.2013.07.002

35. Buitrago R, Depickère S, Bosseno M-FF, et al (2012) Combination of cytochrome b heteroduplex-assay and sequencing for identification of triatomine blood meals. Infect Genet Evol 12:21–7. doi: 10.1016/j.meegid.2011.09.005

36. Noireau F, Flores R, Gutierrez T, Dujardin JP (1997) Detection of sylvatic dark morphs of *Triatoma infestans* in the Bolivian Chaco. Mem??rias do Inst Oswaldo Cruz. doi: 10.1590/S0074-02761997000500003

37. Guarneri AA, Pereira MH, Diotaiuti L (2000) Influence of the blood meal source

on the development of *Triatoma infestans*, *Triatoma brasiliensis*, *Triatoma sordida*, and *Triatoma pseudomaculata* (Heteroptera, Reduviidae). J Med Entomol 37:373–379.

38. Bussacos ACM, Nakayasu ES, Hecht MM, et al (2011) Redundancy of proteins in the salivary glands of *Panstrongylus megistus* secures prolonged procurement for blood meals. J Proteomics. doi: 10.1016/j.jprot.2011.04.028

39. Carrasco HJ, Segovia M, Londoño JC, et al (2014) *Panstrongylus geniculatus* and four other species of triatomine bug involved in the *Trypanosoma cruzi* enzootic cycle: High risk factors for Chagas' disease transmission in the Metropolitan District of Caracas, Venezuela. Parasit Vectors. doi: 10.1186/s13071-014-0602-7

40. Dorn PL, Monroy C, Curtis A (2006) *Triatoma dimidiata* (Latreille, 1811): A review of its diversity across its geographic range and the relationship among populations. Infect Genet Evol. doi: 10.1016/j.meegid.2006.10.001

41. Segovia M, Carrasco HJ, Martínez CEC, et al (2013) Molecular epidemiologic source tracking of orally transmitted Chagas disease, Venezuela. Emerg Infect Dis. doi: 10.3201/eid1907.121576

42. Miles MA, De Souza AA, Povoa M (1981) Chagas' disease in the Amazon basin III. Ecotopes of ten triatomine bug species (Hemiptera: Reduviidae) from the vicinity of Belem, Para State, Brazil. J Med Entomol 18:266–278.

43. Packchanian A (1940) Experimental transmission of *Trypanosoma cruzi* infection in animals by *Triatoma sanguisuga ambigua*. Public Health Rep 55:1526. doi: 10.2307/4583416

44. Kjos SA, Snowden KF, Craig TM, et al (2008) Distribution and characterization of

canine Chagas disease in Texas. Vet Parasitoloty 152:249–256.

45. Brown EL, Roellig DM, Gompper ME, et al (2010) Seroprevalence of *Trypanosoma cruzi* among eleven potential reservoir species from six states across the southern United States. Vector-Borne Zoonotic Dis 10:757–763. doi: 10.1089/vbz.2009.0009

46. Wood SF (1941) Notes on the distribution and habits of Reduviid vectors of Chagas' disease in the southwestern United States (Hemiptera, Reduviidae). Pacific Coast Entomological Society

47. Yaeger RG (1988) The prevalence of *Trypanosoma cruzi* infection in armadillos collected at a site near New Orleans, Louisiana. Am J Trop Med Hyg 38:323–6.

48. Hall C a, Polizzi C, Yabsley MJ, Norton TM (2007) *Trypanosoma cruzi* prevalence and epidemiologic trends in lemurs on St. Catherines Island, Georgia. J Parasitol 93:93–6. doi: 10.1645/GE-936R.1

49. Dorn PL, Daigle ME, Combe CL, et al (2012) Low prevalence of Chagas parasite infection in a nonhuman primate colony in Louisiana. J Am Assoc Lab Anim Sci 51:443–447(5).

50. Seibold HR, Wolf RH (1970) American trypanosomiasis (Chagas' disease) in *Hylobates pileatus*. Lab Anim Care 20:514–7.

51. Patterson JS, Gaunt MW (2010) Phylogenetic multi-locus codon models and molecular clocks reveal the monophyly of haematophagous reduviid bugs and their evolution at the formation of South America. Mol Phylogenet Evol 56:608–621. doi: 10.1016/j.ympev.2010.04.038

52. Dorn PL, Perniciaro L, Yabsley MJ, et al (2007) Autochthonous transmission of

Trypanosoma cruzi, Louisiana. Emerg Infect Dis 13:605–7. doi: 10.3201/eid1304.061002

53. Lyman DF, Monteiro F a, Escalante a a, et al (1999) Mitochondrial DNA sequence variation among triatomine vectors of Chagas' disease. Am J Trop Med Hyg 60:377–386.

54. Brisse S, Verhoef J, Tibayrenc M (2001) Characterisation of large and small subunit rRNA and mini-exon genes further supports the distinction of six *Trypanosoma cruzi* lineages. Int J Parasitol 31:1218–1226. doi: 10.1016/S0020-7519(01)00238-7

55. Boakye DA, Tang J, Truc P, et al (1999) Identification of bloodmeals in haematophagous Diptera by cytochrome B heteroduplex analysis. Med Vet Entomol 13:282–287. doi: 10.1046/j.1365-2915.1999.00193.x

56. Vossen RHAM, Aten E, Roos A, Den Dunnen JT (2009) High-resolution melting analysis (HRMA) - More than just sequence variant screening. Hum Mutat 30:860–866. doi: 10.1002/humu.21019

57. BioRad (2012) Precision Melt Analysis ™ Software.

58. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. doi: 10.1093/molbev/msw054

59. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452. doi: 10.1093/bioinformatics/btp187

60. Bandelt H (1994) Phylogenetic networks. Verh. Naturwiss. Ver. Hambg

61. data.lsuinsects.org (2017). See Appendix D for full citation.

62. http://www.idigbio.org/portal (2017). See Appendix D for full citation.

63. www.iNaturalist.org/observations (2017). See Appendix D for full citation.

64. (2017) Symbiota Collections of Arthropods Network.

65. (2016) census.gov. In: TIGER Geod.

66. Phillips S, Dudík M, Schapire R (2004) A maximum entropy approach to species distribution modeling. Proc Twenty-First Int Conf Mach Learn 655–662. doi: 10.1145/1015330.1015412

67. Brown JL (2014) SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol Evol 5:694–700. doi: 10.1111/2041-210X.12200

68. Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography (Cop) 33:607–611. doi: 10.1111/j.1600-0587.2009.06142.x

69. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–26. doi: 10.1093/molbev/msl149

70. Gomez-Palacio A, Arboleda S, Dumonteil E, et al (2015) Ecological niche and geographic distribution of the Chagas disease vector, *Triatoma dimidiata* (Reduviidae: Triatominae): Evidence for niche differentiation among cryptic species. Infect Genet Evol 36:15–22. doi: 10.1016/j.meegid.2015.08.035

71. Pfeiler E, Bitler BG, Ramsey JM, et al (2006) Genetic variation, population structure, and phylogenetic relationships of *Triatoma rubida* and *T. recurva*

(Hemiptera: Reduviidae: Triatominae) from the Sonoran Desert, insect vectors of the Chagas' disease parasite *Trypanosoma cruzi*. Mol Phylogenet Evol 41:209–221. doi: 10.1016/j.ympev.2006.07.001

72. Monteiro FA, Donnelly MJ, Beard CB, Costa J (2004) Nested clade and phylogeographic analyses of the Chagas disease vector *Triatoma brasiliensis* in northeast Brazil. Mol Phylogenet Evol 32:46–56. doi: 10.1016/j.ympev.2003.12.011

73. Martinez FH, Villalobos GC, Cevallos AM, et al (2006) Taxonomic study of the Phyllosoma complex and other triatomine (Insecta: Hemiptera: Reduviidae) species of epidemiological importance in the transmission of Chagas disease: Using ITS-2 and mtCytB sequences. Mol Phylogenet Evol 41:279–287. doi: 10.1016/j.ympev.2006.05.002

74. Dumonteil E, Tripet F, Ramirez-Sierra MJ, et al (2007) Assessment of *Triatoma dimidiata* dispersal in the Yucatan Peninsula of Mexico by morphometry and microsatellite markers. Am J Trop Med Hyg 76:930–937.

75. Luca F, Hudson RR, Witonsky DB, et al (2011) A reduced representation approach to population genetic analyses and applications to human evolution. Genome Res 21:1087–1098. doi: 10.1101/gr.119792.110

76. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48. doi: 10.1093/oxfordjournals.molbev.a026036

77. Roux E, Venâncio ADF, Girres J, et al (2011) Spatial patterns and eco-epidemiological systems – part II : characterising spatial patterns of the occurrence

of the insect vectors of Chagas disease based on remote sensing and field data. Heal (San Fr 6:53–64.

78. Suarez-Davalos V, Dangles O, Villacis AG, Grijalva MJ (2010) Microdistribution of sylvatic Triatomine populations in central-coastal Ecuador. J Med Entomol 47:80–88. doi: 10.1603/033.047.0111

79. Grijalva MJ, Terán D, Dangles O (2014) Dynamics of sylvatic Chagas disease vectors in coastal Ecuador is driven by changes in land cover. PLoS Negl Trop Dis. doi: 10.1371/journal.pntd.0002960

80. Aslett M, Aurrecoechea C, Berriman M, et al (2009) TriTrypDB: A functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. doi: 10.1093/nar/gkp851

81. Breniere SF, Waleckx E, Barnabe C (2016) Over six thousand *Trypanosoma cruzi* strains classified into discrete typing units (DTUs): Attempt at an inventory. PLoS Negl Trop Dis 10:1–19. doi: 10.1371/journal.pntd.0004792

82. Curtis-Robles R (2016) Ecoepidemiology of *Trypanosoma cruzi* in Texas. Texas A&M University

83. Curtis-Robles R, Lewis BC, Hamer SA (2016) High *Trypanosoma cruzi* infection prevalence associated with minimal cardiac pathology among wild carnivores in central Texas. Int J Parasitol Parasites Wildl 5:117–123. doi: 10.1016/j.ijppaw.2016.04.001

84. Ramírez JD, Guhl F, Rendón LM, et al (2010) Chagas cardiomyopathy manifestations and *Trypanosoma cruzi* genotypes circulating in chronic chagasic patients. PLoS Negl Trop Dis 4:1–9. doi: 10.1371/journal.pntd.0000899

85. Monteiro WM, Magalhães LKC, de Sá ARN, et al (2012) *Trypanosoma cruzi* IV causing outbreaks of acute Chagas disease and infections by different haplotypes in the western Brazilian Amazonia. PLoS One. doi: 10.1371/journal.pone.0041284

86. Lowery GHJ (1974) The mammals of Louisiana and its adjacent waters. Louisiana State University Press, Nashville

87. Shender L, Niemela M, Conrad P, et al (2016) Habitat management to reduce human exposure to *Trypanosoma cruzi* and western conenose bugs (*Triatoma protracta*). Ecohealth 13:525–534. doi: 10.1007/s10393-016-1153-5

88. Rowland ME, Maloney J, Cohen S, et al (2010) Factors associated with *Trypanosoma cruzi* exposure among domestic canines in Tennessee. J Parasitol 96:547–551. doi: 10.1645/GE-2299.1

89. Barr S, Dennis V, Klei T (1991) Serologic and blood culture survey of *Trypanosoma cruzi* infection in four canine populations in southern Louisiana. Am J Vet Res 52:570–573.

90. Gürtler RE, Cardinal MV (2015) Reservoir host competence and the role of domestic and commensal hosts in the transmission of *Trypanosoma cruzi*. Acta Trop. doi: 10.1016/j.actatropica.2015.05.029

91. Gürtler RE, Solard ND, Lauricela MA, et al (1986) Dynamics of transmission of *Trypanosoma cruzi* in a rural area of Argentina. III. Persistence of *T. cruzi* parasitemia among canine reservoirs in a two-year follow-up. Rev Inst Med Trop Sao Paulo 28:213–219.

92. Gürtler RE, Cécere MC, Rubel DN, et al (1991) Chagas disease in north-west Argentina: Infected dogs as a risk factor for the domestic transmission of

Trypanosoma cruzi. Trans R Soc Trop Med Hyg 85:741–745. doi: 10.1016/0035-9203(91)90440-A

93. Charles R a, Kjos S, Ellis AE, et al (2013) Southern plains woodrats (*Neotoma micropus*) from southern Texas are important reservoirs of two genotypes of *Trypanosoma cruzi* and host of a putative novel *Trypanosoma* species. Vector Borne Zoonotic Dis 13:22–30. doi: 10.1089/vbz.2011.0817
94. Farghadani R, Babadi AA (2015) Nucleotide variation of the mitochondrial cytochrome B gene in the Malay population. Rom J Leg Med 23:57–60. doi: 10.4323/rjlm.2015.57
95. Amer SAM, Alhothali B, Tubaigy SMA (2015) Possible application of cytochrome b gene for human identification. Arab J Forensic Sci Forensic Med 1:24–248. doi: 10.12816/0017705
96. Bataille M, Crainic K, Leterreux M, et al (1999) Multiplex amplification of mitochondrial DNA for human and species identification in forensic evaluation. Forensic Sci Int 99:165–70. doi: 10.1016/S0379-0738(98)00185-6
97. Louisiana State Arthropod Museum.
98. Merow C, Smith MJ, Silander JA (2013) A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. Ecography (Cop) 36:1058–1069. doi: 10.1111/j.1600-0587.2013.07872.x
99. (2015) Chagas Disease - Biology. In: Cent. Dis. Control Prev. - Div. Parasit. Dis. https://www.cdc.gov/parasites/chagas/biology.html. Accessed 3 Apr 2017
100. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425. doi: citeulike-article-

id:93683

101. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution (N Y) 39:783–791. doi: 10.2307/2408678

Appendix A

>Haplotype 1
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATACTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
TAAATTAGGGGGAGTTATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 2
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCGTTAGTAACACCAGTGC
ACATTCAACCAGAATGATACTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
TAAATTAGGGGGAGTCATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 3
CGATTCTAATAGATCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATATTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
TAAATTAGGGGGAGTTATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 4
CGATTCTAATAGATCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATATTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
TAAATTAGGGGGAGTTATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAT
>Haplotype 5
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATATTTCCTATTTGCATACGCAATCTTACGATCCATTCCTAA
TAAATTAGGGGGAGTTATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTCCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 6
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATACTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
CAAATTAGGGGGAGTCATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 7
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATATTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
TAAATTAGGGGGAGTTATTGCAATAGTCTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 8
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATACTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
CAAATTAGGGGGAGTCATTGCAATAGTTTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA
>Haplotype 9
CGATTCTAATAGACCCAGAAAACTTTATTCCTGCAAACCCATTAGTAACACCAGTGC
ACATTCAACCAGAATGATACTTCCTATTTGCATACGCAATTTTACGATCCATTCCTAA
CAAATTAGGGGGAGTCATTGCAATAGTTTCATCAATTGCAATTATTTTAATCCTTCCA
TTTACTAATAAAAGAAAATTTCAAGGCCTCCCAA

Appendix B

Full Citations for Collection Data:

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"exists": {"field": "geopoint"}}, {"term": {"scientificname": "triatoma sanguisuga"}}]}}},
137 records, accessed on 2017-03-04T18:01:41.003398,
contributed by 3 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/5975bbda-cd92-4084-8a09-ce1e28e6164f (133 records)
http://www.idigbio.org/portal/recordsets/ff111763-e72d-4f24-8914-b5b2dd94908c (3 records)
http://www.idigbio.org/portal/recordsets/7fcdca8e-7469-480c-8516-cce4e24c37c9 (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "sigmodon hispidus"}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "sigmodon hispidus"}}}}]}}},
15083 records, accessed on 2017-02-27T19:10:49.167779,
contributed by 43 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (3113 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (2956 records)
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (1883 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (1853 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (1273 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (867 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (432 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (419 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (392 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (235 records)
http://www.idigbio.org/portal/recordsets/a6eee223-cf3b-4079-8bb2-b77dad8cae9d (225 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (212 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (206 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (198 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (158 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (115 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (112 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (72 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (56 records)
http://www.idigbio.org/portal/recordsets/e165b318-d5f7-40d5-a0d9-82ba3c31060f (49 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (45 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (34 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (33 records)
http://www.idigbio.org/portal/recordsets/9ab47b07-99a9-4509-884b-be9383908b28 (28 records)
http://www.idigbio.org/portal/recordsets/9e046dad-2b23-4f95-8eaf-c0346de2556e (24 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (23 records)
http://www.idigbio.org/portal/recordsets/271a9ce9-c6d3-4b63-a722-cb0adc48863f (15 records)
http://www.idigbio.org/portal/recordsets/cd9e44d9-40cd-4a75-ab9c-4db2ac9e87ea (10 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (9 records)
http://www.idigbio.org/portal/recordsets/79dfdec6-3e24-489c-a7ce-85dcc52bc3f9 (8 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (6 records)
http://www.idigbio.org/portal/recordsets/cb790bee-26da-40ed-94e0-d179618f9bd4 (4 records)
http://www.idigbio.org/portal/recordsets/319636db-c2da-493c-beac-1194949e95b4 (3 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (3 records)
http://www.idigbio.org/portal/recordsets/fb4a4330-9124-4013-a0e1-af42ee20cd16 (2 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (2 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (2 records)
http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (1 records)
http://www.idigbio.org/portal/recordsets/fcbcb214-cd62-4453-af56-b4b49161a261 (1 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (1 records)
http://www.idigbio.org/portal/recordsets/0f3f26e2-cc13-47a3-a268-4c321b621586 (1 records)
http://www.idigbio.org/portal/recordsets/a5fe6f13-2121-41dd-a036-dae15546ad91 (1 records)
http://www.idigbio.org/portal/recordsets/929bf047-9ad7-48bd-88fa-c2630d423e8a (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "sciurus niger"}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "sciurus niger"}}}}]}}},
2500 records, accessed on 2017-02-27T19:03:54.374900,
contributed by 36 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (918 records)

http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (337 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (202 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (145 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (130 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (121 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (99 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (61 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (53 records)
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (48 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (44 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (43 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (35 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (31 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (27 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (27 records)
http://www.idigbio.org/portal/recordsets/da67ebd9-52de-444d-b114-e23c03111ac6 (26 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (22 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (19 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (18 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (14 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (13 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (13 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (9 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (8 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (8 records)
http://www.idigbio.org/portal/recordsets/fcbcb214-cd62-4453-af56-b4b49161a261 (5 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (4 records)
http://www.idigbio.org/portal/recordsets/2d94a3ac-f505-49ec-98e7-3b7dc48344dd (4 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (4 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (3 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (3 records)
http://www.idigbio.org/portal/recordsets/e165b318-d5f7-40d5-a0d9-82ba3c31060f (2 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (2 records)
http://www.idigbio.org/portal/recordsets/9e046dad-2b23-4f95-8eaf-c0346de2556e (1 records)
http://www.idigbio.org/portal/recordsets/0f3f26e2-cc13-47a3-a268-4c321b621586 (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "sciurus carolinensis"}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "sciurus carolinensis "}}}}]}}},
2758 records, accessed on 2017-02-27T19:02:33.437360,
contributed by 37 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (883 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (407 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (319 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (273 records)
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (118 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (117 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (107 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (80 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (62 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (50 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (36 records)
http://www.idigbio.org/portal/recordsets/da67ebd9-52de-444d-b114-e23c03111ac6 (32 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (31 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (29 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (25 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (21 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (21 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (21 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (16 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (16 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (15 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (14 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (11 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (9 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (7 records)
http://www.idigbio.org/portal/recordsets/a5fe6f13-2121-41dd-a036-dae15546ad91 (7 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (6 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (5 records)
http://www.idigbio.org/portal/recordsets/2d94a3ac-f505-49ec-98e7-3b7dc48344dd (4 records)

http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (3 records)
http://www.idigbio.org/portal/recordsets/76015dea-c909-4e6d-a8e1-3bf35763571e (3 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (2 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (2 records)
http://www.idigbio.org/portal/recordsets/311c3a01-c824-4a85-8771-fcc3f353619b (2 records)
http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (2 records)
http://www.idigbio.org/portal/recordsets/9e046dad-2b23-4f95-8eaf-c0346de2556e (1 records)
http://www.idigbio.org/portal/recordsets/271a9ce9-c6d3-4b63-a722-cb0adc48863f (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"terms": {"execution": "or", "scientificname": ["procyon lotor", ""]}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "procyon lotor"}}}}]}}},
3700 records, accessed on 2017-02-27T18:12:33.763898,
contributed by 38 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (1109 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (406 records)
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (374 records)
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (366 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (292 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (138 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (124 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (115 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (109 records)
http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (89 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (56 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (55 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (53 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (52 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (43 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (40 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (39 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (37 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (32 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (30 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (21 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (20 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (20 records)
http://www.idigbio.org/portal/recordsets/da67ebd9-52de-444d-b114-e23c03111ac6 (16 records)
http://www.idigbio.org/portal/recordsets/76015dea-c909-4e6d-a8e1-3bf35763571e (10 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (8 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (7 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (6 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (6 records)
http://www.idigbio.org/portal/recordsets/fb4a4330-9124-4013-a0e1-af42ee20cd16 (5 records)
http://www.idigbio.org/portal/recordsets/b3d3a357-9fa6-453c-9f02-d86a1bbc762a (5 records)
http://www.idigbio.org/portal/recordsets/a5fe6f13-2121-41dd-a036-dae15546ad91 (4 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (3 records)
http://www.idigbio.org/portal/recordsets/0f3f26e2-cc13-47a3-a268-4c321b621586 (3 records)
http://www.idigbio.org/portal/recordsets/e165b318-d5f7-40d5-a0d9-82ba3c31060f (2 records)
http://www.idigbio.org/portal/recordsets/319636db-c2da-493c-beac-1194949e95b4 (2 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (2 records)
http://www.idigbio.org/portal/recordsets/9e046dad-2b23-4f95-8eaf-c0346de2556e (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "peromyscus gossypinus"}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "peromyscus gossypinus"}}}}]}}},
3785 records, accessed on 2017-02-27T19:05:28.909169,
contributed by 26 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (1319 records)
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (734 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (366 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (356 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (274 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (168 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (137 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (83 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (80 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (80 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (71 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (33 records)

http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (21 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (15 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (12 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (7 records)
http://www.idigbio.org/portal/recordsets/271a9ce9-c6d3-4b63-a722-cb0adc48863f (7 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (5 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (4 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (4 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (3 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (2 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (1 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (1 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (1 records)
http://www.idigbio.org/portal/recordsets/79dfdec6-3e24-489c-a7ce-85dcc52bc3f9 (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "neotoma floridana"}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "neotoma floridana"}}}}]}}},
1262 records, accessed on 2017-02-27T18:09:03.566956,
contributed by 30 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (274 records)
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (255 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (251 records)
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (236 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (34 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (31 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (25 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (21 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (19 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (14 records)
http://www.idigbio.org/portal/recordsets/271a9ce9-c6d3-4b63-a722-cb0adc48863f (14 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (13 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (9 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (8 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (7 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (7 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (7 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (6 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (5 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (5 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (5 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (3 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (2 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (2 records)
http://www.idigbio.org/portal/recordsets/fcbcb214-cd62-4453-af56-b4b49161a261 (2 records)
http://www.idigbio.org/portal/recordsets/79dfdec6-3e24-489c-a7ce-85dcc52bc3f9 (2 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (2 records)
http://www.idigbio.org/portal/recordsets/cb790bee-26da-40ed-94e0-d179618f9bd4 (1 records)
http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (1 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "mephitis mephitis"}}, {"exists": {"field": "geopoint"}}, {"query": {"match": {"_all": {"operator": "and", "query": "mephitis mephitis"}}}}]}}},
1555 records, accessed on 2017-02-27T20:27:14.223832,
contributed by 36 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (531 records)
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (274 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (167 records)
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (64 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (61 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (51 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (46 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (43 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (38 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (36 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (33 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (32 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (26 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (25 records)

http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (23 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (16 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (11 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (9 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (9 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (8 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (8 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (7 records)
http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (6 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (4 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (4 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (3 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (3 records)
http://www.idigbio.org/portal/recordsets/0f3f26e2-cc13-47a3-a268-4c321b621586 (3 records)
http://www.idigbio.org/portal/recordsets/cb790bee-26da-40ed-94e0-d179618f9bd4 (3 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (3 records)
http://www.idigbio.org/portal/recordsets/e165b318-d5f7-40d5-a0d9-82ba3c31060f (2 records)
http://www.idigbio.org/portal/recordsets/a5fe6f13-2121-41dd-a036-dae15546ad91 (2 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (1 records)
http://www.idigbio.org/portal/recordsets/2d94a3ac-f505-49ec-98e7-3b7dc48344dd (1 records)
http://www.idigbio.org/portal/recordsets/fcbcb214-cd62-4453-af56-b4b49161a261 (1 records)
http://www.idigbio.org/portal/recordsets/c530ad19-9847-4ea7-a807-f6753c3936d6 (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "didelphis virginiana"}}, {"exists": {"field": "geopoint"}},
{"query": {"match": {"_all": {"operator": "and", "query": "didelphis virginiana"}}}}]}}},
2491 records, accessed on 2017-02-27T20:24:10.652282,
contributed by 37 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/f9a33279-d6ba-41c7-a511-ef6adfcb6e20 (850 records)
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (372 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (245 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (159 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (142 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (137 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (78 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (62 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (57 records)
http://www.idigbio.org/portal/recordsets/50cfe20a-9100-4710-89f9-a97bc3aa53d7 (49 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (48 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (45 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (33 records)
http://www.idigbio.org/portal/recordsets/5a262e5d-0605-4067-ba71-3fd578c3c6bb (32 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (22 records)
http://www.idigbio.org/portal/recordsets/76015dea-c909-4e6d-a8e1-3bf35763571e (18 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (17 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (16 records)
http://www.idigbio.org/portal/recordsets/62c35d43-f15c-451d-a8be-1b9c6928b8bd (14 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (13 records)
http://www.idigbio.org/portal/recordsets/a5fe6f13-2121-41dd-a036-dae15546ad91 (12 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (10 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (10 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (8 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (8 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (7 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (6 records)
http://www.idigbio.org/portal/recordsets/e165b318-d5f7-40d5-a0d9-82ba3c31060f (5 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (4 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (3 records)
http://www.idigbio.org/portal/recordsets/a6eee223-cf3b-4079-8bb2-b77dad8cae9d (2 records)
http://www.idigbio.org/portal/recordsets/0e2f3962-e905-48f2-a1c6-19d16e2bd5ba (2 records)
http://www.idigbio.org/portal/recordsets/2d94a3ac-f505-49ec-98e7-3b7dc48344dd (1 records)
http://www.idigbio.org/portal/recordsets/319636db-c2da-493c-beac-1194949e95b4 (1 records)
http://www.idigbio.org/portal/recordsets/9e046dad-2b23-4f95-8eaf-c0346de2556e (1 records)
http://www.idigbio.org/portal/recordsets/fcbcb214-cd62-4453-af56-b4b49161a261 (1 records)
http://www.idigbio.org/portal/recordsets/9ab47b07-99a9-4509-884b-be9383908b28 (1 records)

http://www.idigbio.org/portal (2017),
Query: {"filtered": {"filter": {"and": [{"term": {"scientificname": "dasypus novemcinctus"}}, {"exists": {"field": "geopoint"}},
{"query": {"match": {"_all": {"operator": "and", "query": "dasypus novemcinctus"}}}}]}}},
562 records, accessed on 2017-02-27T19:07:45.139618,

contributed by 40 Recordsets, Recordset identifiers:
http://www.idigbio.org/portal/recordsets/9e046dad-2b23-4f95-8eaf-c0346de2556e (100 records)
http://www.idigbio.org/portal/recordsets/c6969e30-ca21-4576-954d-9c0e052bdde9 (78 records)
http://www.idigbio.org/portal/recordsets/c2dcb184-6c90-4aa3-9ebb-33b2d53837b9 (78 records)
http://www.idigbio.org/portal/recordsets/48e1b8c1-91aa-4b87-8ca0-de1f81232eaf (47 records)
http://www.idigbio.org/portal/recordsets/0a0f5c81-bf4d-492b-b459-08bd987a0c9a (40 records)
http://www.idigbio.org/portal/recordsets/d0105f1d-a9a0-4cd4-817d-aebfb5512923 (37 records)
http://www.idigbio.org/portal/recordsets/a6eee223-cf3b-4079-8bb2-b77dad8cae9d (34 records)
http://www.idigbio.org/portal/recordsets/ec359278-df8e-4766-a1d3-4b55fd822704 (17 records)
http://www.idigbio.org/portal/recordsets/a5fe6f13-2121-41dd-a036-dae15546ad91 (14 records)
http://www.idigbio.org/portal/recordsets/76015dea-c909-4e6d-a8e1-3bf35763571e (13 records)
http://www.idigbio.org/portal/recordsets/09b18522-5643-478f-86e9-d2e34440d43e (13 records)
http://www.idigbio.org/portal/recordsets/433646ab-571a-44f5-820e-25e0736b1113 (10 records)
http://www.idigbio.org/portal/recordsets/6e6e2b47-fa3e-4bd9-8f1c-105b741d31df (10 records)
http://www.idigbio.org/portal/recordsets/b4d4e884-a2ef-4967-b4cb-2072fc465eaf (8 records)
http://www.idigbio.org/portal/recordsets/f3d1fbbb-93d5-432e-8808-ebc08c42ef6d (7 records)
http://www.idigbio.org/portal/recordsets/2ee6534e-46ab-4233-98c4-d13c4262ce2e (6 records)
http://www.idigbio.org/portal/recordsets/14a8f79f-eab7-48da-ad50-bda142703820 (5 records)
http://www.idigbio.org/portal/recordsets/552ce2e5-b627-4d6d-b914-6b495d0a79e6 (4 records)
http://www.idigbio.org/portal/recordsets/21ea1ef6-4a2a-4ff2-a18b-5c0f297fc1cf (4 records)
http://www.idigbio.org/portal/recordsets/9dce915b-3de4-4a7d-a68d-e4c4c15809ce (3 records)
http://www.idigbio.org/portal/recordsets/2d94a3ac-f505-49ec-98e7-3b7dc48344dd (3 records)
http://www.idigbio.org/portal/recordsets/9d8ced48-62c5-4ce0-99e7-a03550c674c0 (3 records)
http://www.idigbio.org/portal/recordsets/4c512712-3b53-4403-a059-d31d44c7d62d (3 records)
http://www.idigbio.org/portal/recordsets/d9d38b3f-5173-4051-98a6-2efad16fc8da (3 records)
http://www.idigbio.org/portal/recordsets/fb4a4330-9124-4013-a0e1-af42ee20cd16 (2 records)
http://www.idigbio.org/portal/recordsets/93341fe7-38f8-4ef2-8dfc-ae550aa522dc (2 records)
http://www.idigbio.org/portal/recordsets/da67ebd9-52de-444d-b114-e23c03111ac6 (2 records)
http://www.idigbio.org/portal/recordsets/b5d8168e-c310-4870-aa88-eeb3c25256fd (2 records)
http://www.idigbio.org/portal/recordsets/e6d3c1da-a02f-43a2-a5ef-6a035298b933 (2 records)
http://www.idigbio.org/portal/recordsets/9ab47b07-99a9-4509-884b-be9383908b28 (2 records)
http://www.idigbio.org/portal/recordsets/c9316f11-d955-4472-a276-6a26a6514590 (1 records)
http://www.idigbio.org/portal/recordsets/86b1f54d-ac01-4c5e-8ed8-09da2689c7a9 (1 records)
http://www.idigbio.org/portal/recordsets/e165b318-d5f7-40d5-a0d9-82ba3c31060f (1 records)
http://www.idigbio.org/portal/recordsets/fcbcb214-cd62-4453-af56-b4b49161a261 (1 records)
http://www.idigbio.org/portal/recordsets/17969b7f-c1d0-4c84-9cbc-de64b90a62a5 (1 records)
http://www.idigbio.org/portal/recordsets/5c861676-8285-4a04-b1c5-94ce73342320 (1 records)
http://www.idigbio.org/portal/recordsets/cb790bee-26da-40ed-94e0-d179618f9bd4 (1 records)
http://www.idigbio.org/portal/recordsets/271a9ce9-c6d3-4b63-a722-cb0adc48863f (1 records)
http://www.idigbio.org/portal/recordsets/929bf047-9ad7-48bd-88fa-c2630d423e8a (1 records)
http://www.idigbio.org/portal/recordsets/adae5c6c-72f3-4cd8-a00b-3ea71d516abc (1 records)

data.lsuinsects.org (2017),
Query: “genus” = “Triatoma”; “species” = sanguisuga
137 records, accessed on 2017-3-4T17:43
LSU:LSAM:119425 – LSU:LSAM:119449
LSU:LSAM:153164 – LSU:LSAM:153204
LSU:LSAM:153206
LSU:LSAM:153208 – LSU:LSAM:153218
LSU:LSAM:153220 – LSU:LSAM:153236
LSU:LSAM:157228
LSU:LSAM:157230 – LSU:LSAM:157231
LSU:LSAM:157233 – LSU:LSAM:157236
LSU:LSAM:157293 – LSU:LSAM:157326
LSU:LSAM:278235

www.iNaturalist.org/observations (2017),
Query: q=triatoma+sanguisuga&search_on=names&quality_grade=any&identifications=any&taxon_id=322394&verifiable=true
Columns id, url, place_guess, latitude, longitude, positional_accuracy, geoprivacy, positioning_method, positioning_device, species_guess, scientific_name, common_name, iconic_taxon_name, taxon_id
22 records, accessed on 2017-3-4T18:03
http://www.inaturalist.org/observations/333799
http://www.inaturalist.org/observations/579731
http://www.inaturalist.org/observations/745800
http://www.inaturalist.org/observations/1216537
http://www.inaturalist.org/observations/1690537
http://www.inaturalist.org/observations/1767523
http://www.inaturalist.org/observations/1935668
http://www.inaturalist.org/observations/2679849
http://www.inaturalist.org/observations/3045781

http://www.inaturalist.org/observations/3188037
http://www.inaturalist.org/observations/3243381
http://www.inaturalist.org/observations/3463785
http://www.inaturalist.org/observations/3507808
http://www.inaturalist.org/observations/3670166
http://www.inaturalist.org/observations/3757383
http://www.inaturalist.org/observations/3785183
http://www.inaturalist.org/observations/3840164
http://www.inaturalist.org/observations/3869895
http://www.inaturalist.org/observations/3948532
http://www.inaturalist.org/observations/4223075
http://www.inaturalist.org/observations/4330854
https://www.inaturalist.org/observations/5213600