2019

Protein Mass Spectrometry Aids In Chagas Vector Blood Meal Identification And Offers An Innovative Approach To Battling Vector-Borne Diseases

Judith Ina Keller
University of Vermont

PROTEIN MASS SPECTROMETRY AIDS IN CHAGAS VECTOR BLOOD MEAL IDENTIFICATION AND OFFERS AN INNOVATIVE APPROACH TO BATTLING VECTOR-BORNE DISEASES

A Dissertation Presented

by

Judith Ina Keller

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Biology

January, 2019

Defense Date: October 26, 2018
Dissertation Examination Committee:

Lori Stevens, Ph.D., Advisor
Bryan Ballif, Ph.D., Advisor
Jason Botten, Ph.D., Chairperson
Melissa Pespeni, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

Vector borne-diseases make up a significant portion of morbidity and mortality worldwide, being responsible for around 700,000 deaths annually according to the World Health Organization. Neglected, tropical diseases such as Chagas disease have a significant impact on people in Latin America, affecting millions, and especially those residing in rural areas. Chagas disease is the number one cause for heart disease in Latin America, and is caused by the *Trypanosoma cruzi* parasite, carried by Triatominae insect vectors. The intricate life cycle of the parasite, ecology and behavior of the vector, and lack of disease treatment options, make Chagas disease challenging to control. Prevention measures are highly sought after, and implementation science approaches such as Ecohealth management engage affected communities in disease prevention. Knowing what insect vectors are feeding on sheds light on vector ecology and behavior, aiding in vector management which is pivotal in disease prevention.

While DNA-based methods have traditionally been used to study vector blood meals, they come with limitations and challenges, such as the need for fresh, high abundance blood meals. Therefore, the goal of this research was to evaluate Chagas vector blood meal sources using an innovative protein mass spectrometry-based approach. We demonstrate first the ability to utilize liquid chromatography tandem mass spectrometry (LC-MS/MS) to correctly identify hemoglobin protein peptides from mouse blood and subsequently identify Chagas vector blood meal sources from field-collected insect vectors where blood meal identification is compared with traditional DNA-based methods as a control.

An experimental feeding study allowed us to then demonstrate the longevity of hemoglobin protein peptides for blood meal detection, showing LC-MS/MS-based blood meal identification outperforms DNA-based polymerase chain reaction (PCR) at least 4 weeks post-feeding and 12 weeks post-molting. This allowed us to test the limits of our innovative detection method experimentally and comparatively.

Finally, we evaluated blood meals in field-caught insect vectors collected as part of a large collaborative Ecohealth project in Central America. LC-MS/MS identified two times as many blood meals in insect vectors, including those that did not have blood meals detected with DNA-based PCR. As single vectors often feed on multiple sources, we also validated our ability to decipher multiple blood meals from an individual vector and showed the ability to quantify a blood meal using synthetic AQUA (Absolute QUAntification) peptides, a first step in using quantification data for identifying blood meals not currently in our underlying database. Furthermore, we show that lower resolution mass spectrometers are able to identify blood meals from taxa correctly, an important and strong attribute of our LC-MS/MS-based method, opening the door to using proteomics in countries where Chagas disease is endemic and resources are limited.

Even though expertise and resources of research labs differ in locations across the globe, herein is described how LC-MS/MS is a valuable additional tool for fighting neglected tropical diseases. Ultimately, hemoglobin-based LC-MS/MS vector blood meal identification is a complementary technique to available molecular methods and can confidently identify Chagas vector blood meal sources to aid in understanding vector biology and ecology, and aid in developing appropriate Ecohealth vector control measures.

CITATIONS

Material from this dissertation has been published in the following form:

Keller J.I., Ballif B.A., St. Clair R.M., Vincent J.J., Monroy M.C., Stevens L.. 2017. Chagas disease vector blood meal sources identified by protein mass spectrometry. PLOS ONE. 12(12):e0189647.

Keller J.I., Schmidt J.O., Schmoker A.M., Ballif B.A., Stevens L.. 2018. Protein mass spectrometry extends temporal blood meal detection over polymerase chain reaction in mouse-fed Chagas disease vectors. Memórias do Instituto Oswaldo Cruz. 113(10).

EPIGRAPH

"We have a choice to use the gift of our lives to make the world a better place."
Jane Goodall

DEDICATION

To Lucy,

"Let me tell you the secret that has led me to my goal.

My strength lies solely in my tenacity." -Louis Pasteur

and

Justin

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisors, Dr. Bryan Ballif and Dr. Lori Stevens, for their constant advice, support, and encouragement throughout my time at UVM. Thank you to Bryan for always being ready to jump in with "negative controls" he found in the hallway and all the advice on everything mass spectrometry. Thank you to Lori for always taking time to chat about "lab safety" and all the time she spent working with me to improve my writing. Thank you both for your valuable input, support, and investment in me and in this project. Thank you also to my committee, Dr. Melissa Pespeni and Dr. Jason Botten, for helping me think outside of the box and applying my research to the bigger picture. Thank you to all the collaborators on these various projects, especially Dr. Carlota Monroy, Dr. Patricia Dorn, and Dr. Justin Schmidt- this work would not have been possible without you.

Thank you to the Biology department, my fellow grad students and lab members, especially Yainna Hall, Raquel Lima, Joseph Gallant, Anna Schmoker, Dr. Marion Weir, and Riley St. Clair for always being there with listening ears and heartfelt advice. Thank you to all the amazing people in my life that have guided me and shaped me to become a scientist. As this journey comes to an end, thank you to my incredible family for their persistence in support over the many years. Thank you to my wonderful husband Justin for supporting me through the highs and lows. And thank you to my Lucy for being part of this particular adventure from the very beginning.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1: INTRODUCTION

1.1. Vector-borne diseases – a worldwide problem

Vector-borne diseases are responsible for over 17 % of all infectious diseases affecting humans worldwide and cause substantial morbidity and mortality, including 700,000 deaths annually (World Health Organization 2017). Although disease vectors include ticks and snails, most vector-borne diseases are transmitted by insects. Insect vectors are often hematophagous, meaning they ingest blood for sustenance, and have evolved to digest blood efficiently (Mesquita et al. 2015). It is usually during one of these feedings that insect vectors have the potential to transmit different disease-causing, or etiological agents between humans, or between humans and other animals. A multitude of the different vector-borne diseases caused by bacteria, viruses, or parasites disproportionally affect poor regions of the world, particularly in tropical and subtropical climates (World Health Organization 2017). In Latin America, Chagas disease, an insect vector-borne disease caused by the *Trypanosoma cruzi* parasite, which is related to the etiological agent of African Sleeping Sickness, occurs in many rural communities and is considered a neglected tropical disease.

Indeed, Chagas disease has plagued people for at least 9,000 years as was shown by evidence of *T. cruzi* infection in mummies from Chile and Peru (Aufderheide et al. 2004). Today, it is estimated that 6-7 million people are infected with Chagas disease and 10-15 % of people living in Latin America live in areas where they are at risk for infection (World Health Organization 2015). Most commonly, Chagas disease is transmitted by a reduviid insect vector in the Triatominae subfamily. These insect

vectors take a blood meal, and shortly after, defecate on the host they are feeding on. The Chagas parasite-laden feces can then be introduced into the host through a break in the skin or mucous membrane. More common for domestic animals like dogs, the parasite can also be taken up by ingesting vector feces, sometimes when eating whole vectors (Bern et al. 2011).

Inherently challenging to any vector borne disease, is the interaction of the insect vector and the vertebrate hosts the vectors feed on. In the case of Chagas disease, the disease transmission cycle is further complicated given the complex life cycle of the *T. cruzi* parasite. The parasite undergoes several developmental stages not only within the mammalian host, but also in the insect vector. Within the insect vector, the trypanomastigotes taken up during a blood meal of the vector transform into epimastigotes, which reproduce in the vector midgut (Rassi and Marin-Neto 2010). These epimastigotes then migrate further down the reproductive tract of the vector into the hindgut and rectum where they differentiate into metacyclic trypanomastigotes and then subsequently are shed during defecation. Once the metacyclic trypanomastigotes in the vector's feces are introduced into the mammalian host through an open wound or mucous membrane, they soon penetrate nucleated cells where they differentiate into amastigotes and replicate (Rassi and Marin-Neto 2010). Eventually amastigotes transform back into trypanomastigotes and spread to different tissues when infected cells lyse. This signifies the acute phase of the disease where people have symptoms of general malaise and occurs 1-2 weeks after exposure.

As the immune response in the mammalian host develops, the parasite load drops to mostly undetectable levels, but the parasite is not eradicated without treatment. Untreated, the infected host enters the chronic phase of the disease, which is lifelong (Rassi and Marin-Neto 2010). Approximately 12,000 people die each year from complications of the disease, which, in its chronic stages, causes major cardiomyopathies, megaesophagus, and megacolon. Chagas disease is the major cause of heart disease in Latin America.

While horizontal transmission is more rare (1-10% of cases) (Bern et al. 2011), lack of treatment options for the chronic stages of the disease make Chagas disease difficult to manage from a healthcare perspective (Henao-Martínez et al. 2017). Benznidazole and Nifurtimox, the only two approved drugs for Chagas disease come with significant side effects such as significant weight loss, nausea and vomiting, mental diseases (i.e., depression), cognitive impairment, sleep problems such as insomnia, generalized pain, problems with balance, and kidney problems (Castro and Diaz de Toranzo 1988). In addition, injuries to the ovaries and testes have been reported for both drugs, and other serious complications of septicemia (Castro and Diaz de Toranzo 1988, Coura and De Castro 2002). The drugs are also contraindicated for the use in pregnant women and in the elderly or very debilitated individuals (Coura and De Castro 2002). These various side effects are a major challenge when treating people in highly endemic *T. cruzi* regions, given these regions are the lower economic, rural areas of Latin America. Many of the people in these regions are malnutritioned and frequently come in contact with the vector (Rizzo et al. 2003). To date, reducing

contact between vector and host through vector control management has been the most useful approach in decreasing Chagas transmission (World Health Organization 2012).

The management of Chagas disease at a global scale is directed by the World Health Organization's Global vector control response (GVCR) 2017–2030 which has vowed to "provide strategic guidance to countries and development partners for urgent strengthening of vector control as a fundamental approach to preventing disease and responding to outbreaks" (World Health Organization and UNICEF 2017). In implementing this initiative, local collaborations including PAHO, the Pan American division of the WHO, working with local non-government organizations (NGOs) and governments in Latin America have responded at more local scales. In the past, the Central American Initiative (Guatemala, Honduras, El Salvador, Belize, Nicaragua, Costa Rica, Panama) focused on the regional elimination of *Rhodnius prolixus*, which was an introduced species to Central America, found exclusively in domestic habitats in most of that region. The Southern Cone Initiative, a collaboration between Argentina, Bolivia, Brazil, Chile, Paraguay, and Uruguay aimed at disease interruption mainly by the elimination of *Triatoma infestans* in these countries (Dias 2007). These types of initiatives were largely dependent on cooperation between national and international agencies, and aided by the scientific community (Dias 2007).

The highly adaptable vector *Triatoma dimidiata* which inhabits sylvatic, domestic, and peridomestic ecotopes, is currently the most prevalent vector of Chagas disease in Central America. Because of sylvatic populations, this vector has been challenging to eliminate in its native range and therefore, over recent years more

holistic approaches to the disease system and interrupting the transmission cycle have been undertaken. Implementation science promotes the direct impact of scientific research to affected communities, having the ability to influence healthcare decision and policy (Henao-Martínez et al. 2017) and aims at research having meaningful outcomes through direct application to the field. Ecohealth, a type of implementation science, takes a community, or ecosystem approach, to interrupting Chagas disease transmission by directly involving members of affected communities (Monroy et al. 2009). With various hurdles to overcome for battling Chagas disease, such as complex interactions between the vector, host, and parasite, limited disease treatment option, and lack of a vaccine, blocking transmission is a very viable option using an Ecohealth approach.

Determining what vectors are feeding on not only increases our knowledge of vector ecology but has direct impact on making local vector and Ecohealth management decisions. There are over 150 different species of triatomine insect vectors (Justi and Galvão 2017) which are responsible for transmitting the Chagas parasite, *T. cruzi*. As the Triatominae have evolved to be hematophagous, at the very center of this disease lies the blood meal source prevalence of the insect vector. In more sylvatic environments, vectors often feed on rodents, bats, and opossums, even birds, as younger instar vectors are often associated with burrows and nests as they cannot fly (Bustamante et al. 2014, Stevens et al. 2014, De Urioste-Stone et al. 2015, Buitrago et al. 2016). In domestic environments, however, humans, dogs and local livestock as such chickens are common blood meal sources, although the parasite cannot replicate

in avian species and needs a mammalian host to complete its life cycle (Pellecer et al. 2013, Bustamante et al. 2014, Stevens et al. 2014). Infection rates of the vector with *T. cruzi* vary with location and species. In addition, there can be some bias in parasite detection depending on the detection method used (e.g., microscopy versus PCR-based methods), where microscopy generally underestimates *T. cruzi* prevalence (Lardeux et al. 2016). With the impacts Chagas disease has worldwide, feeding patterns, especially on humans, provide valuable information. Knowing vector blood meal sources elucidates vector ecology and behavior, can aid in evaluating Ecohealth management measures, and can help make practical recommendations for blood meal source species practices.

Many different detection methods exist for detecting what hematophagous insect vectors are feeding on and identifying host animal species. Although many of these detection methods are used across different vector borne disease systems, the rest of this chapter will mostly focus on blood meal detection methods commonly used for Chagas disease vectors.

1.2 Blood meal source detection techniques

Vector host blood choice influences disease pathogenesis (Kent 2009) and vector borne disease pose a significant financial burden in countries and regions where government dynamics are changing or poor to begin with (Hotez et al. 2014). Advances in blood meal detection abilities have revolutionized our ability to tackle the Chagas disease burden in a more cost-effective way using Ecohealth and involving community

members and local governments (Lucero et al. 2013, Pellecer et al. 2013). In addition, cutting-edge research has been emerging that microbial gut communities and blood meal prevalence can influence transmission dynamics of *T. cruzi* (Dumonteil et al. 2018).

Historically, blood meals in arthropod insect vectors have often been detected using serological techniques based on group specific antibodies (Kent 2009). Precipitin and antisera tests, although valuable (Rabinovich et al. 2011), have limitations and have now largely been replaced by molecular-based assays that have the ability to give more specificity of a blood meal (Kent 2009).

DNA-based molecular methods can be broken down into the type of DNA used for analysis as outlined by Kent (2009), such as PCR targeting mitochondrial and nuclear DNA, with the addition of genome-level analyses. Mitochondrial DNA, such as cytochrome b-based molecular techniques coupled with cloning and DNA sequencing (Stevens et al. 2012), group-specific primers (Lima-Cordón et al. 2018), heteroduplex analysis (Buitrago et al. 2012), high resolution melting (Peña et al. 2012), or quantitative (real-time) PCR (Lucero et al. 2014) target specific regions of DNA and have many published primer sequences. As these assays are based on a mitochondrial gene, and there are hundreds to thousands of mitochondria per cell, using cytochrome b can increase sensitivity and probability of detection. Mitochondrial markers such as 12 S ribosomal DNA, a molecule with multiple copies in a cell can be used with species or taxa-specific primers, can be coupled with DNA sequencing, and be used with cloning to detect multiple blood meals. Using 12 S PCR amplification and sequencing has

recently been reviewed by Georgieva and colleagues to elucidate new host records based on archival specimens (Georgieva et al. 2017).

A few studies have looked at using short interspersed nuclear elements, or SINE-based PCR primers for blood meal detection in Chagas vectors based on Walker et al. 2003 and Walker et al. 2004 who designed the primers for forensic analysis of blood. SINEs have thousands of transposable element sequences within a genome, because they can be class, order and/or species-specific they provide a powerful way to identify blood meal sources by PCR.

In recent years with the advent of genomics and shorter-read sequencing technologies to reduce cost, different genomics-based techniques have emerged to identify blood meals in insect vectors (e.g., Kieran et al. 2017), with some methods simultaneously evaluating insect vector genomics, parasite burden and strain, and even vector bacterial communities (Kieran et al. 2017, Dumonteil et al. 2018, Orantes et al. 2018).

Although molecular genetic techniques offer specificity and sensitivity in blood meal detection, blood meal detection frequently depends on high-quality DNA from recently fed vectors (Gómez-Díaz and Figuerola 2010). Therefore, temporal resolution using DNA-based methods is influenced by the rate of digestion of blood. As PCR-based methods need intact fragments of at least a hundred base pairs in order to amplify DNA from a blood meal, digested or older blood meals still present challenges. In addition, blood meals consisting of multiple hosts can require further sequencing and cloning steps and are cost-prohibitive in areas where Chagas disease is prevalent. In

terms of genomics, lower resolution or short-read genome sequencing generally only identifies dominant (i.e., most recent and/or most abundant) blood meal sources, yet it simultaneously identifies vector species, microbiome composition and *T. cruzi* infection state (Orantes et al. 2018) and DNA quality needs to be high. Next-generation sequencing approaches such as those of Kieran and colleagues, involve extensive sample processing, including several PCR steps before sequences are ready for Illumina sequencing (Kieran et al. 2017). Dumonteil et al. devised a DNA barcoding technique based on next-generation sequencing for simultaneous identification of insect species, microbiome composition, *T. cruzi* infection state, and vertebrate host feeding choices (Dumonteil et al. 2018), however there detection of up to 7 species in a single insect suggests contamination problems.

Although highly sensitive results can be obtained with genomic-based methods, DNA needs to be of high quality and in high abundance, reference sequences must be available for vertebrate species in a specific study area, and significant care must be taken to reduce the risk of contamination (Kieran et al. 2017, Orantes et al. 2018). Although multiple blood meals have the potential to be detected using genomic-based methods, the ability to create appropriate bioinformatic resources to analyze complex experimental designs is also necessary (Orantes 2017).

As genomic resources have been arising in the field of blood meal identification, proteomic resources have likely been present in Latin America since the beginning of the age of proteomics, but perhaps due to political and economic challenges, cutting-edge research was suppressed (Padrón and Domont 2014). Novel

protein-based methods have emerged over recent years using protein mass spectrometry to identify blood meals of various insect vector species. Global collaborations have introduced mass spectrometry to areas where neglected tropical diseases such as Chagas disease are prevalent. Various mass spectrometry-based methods have been used to identify blood meals in other vector disease systems, such as in those of mosquitos and ticks.

In mosquitos, matrix-assisted, laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) has not only been used to identify mosquito species, but also to simultaneously identify the blood meal these vectors had ingested (Niare et al. 2016). This method has been applied to experimentally fed as well as field-collected mosquito specimens (Yssouf et al. 2014). Although the method is efficient and quick, this technique requires reference spectra of known blood samples to match blood meals to. In addition, as blood meals degrade, spectra no longer offer accurate blood meal species identification, as methods to-date have not used TOF-TOF but relied on the MS1 only.

Önder et al. developed a highly successful tandem spectral library-based approach to identify blood meals in experimentally fed ticks using liquid chromatography tandem mass spectrometry (LC-MS/MS) (Önder et al. 2013, Önder et al. 2014). Blood meals were digested with trypsin and reference libraries of expected spectra from a particular blood source were made. Unknown tick blood meals are then processed in the same manner, and spectra from the unknown blood meal are matched to known spectra in the spectral library. Although this technique can be highly specific,

it also requires spectral libraries to contain known blood samples and Önder et al. (2013) ran into challenges when facing multiple blood meals as often only the top-scoring blood meal could be identified with high confidence.

Wickramasekara et al. and Laskay et al. both applied a LC-MS/MS-based approach to identifying tick blood meals in experimentally fed vectors looking at various blood protein peptides (Wickramasekara et al. 2008, Laskay et al. 2012, Laskay et al. 2013, Song et al. 2015). While structural proteins such as tubulin and actin had too highly conserved sequences between insect and blood host, transferrins and immunoglobulins only offered limited identification of the blood meal. Albumin and hemoglobin peptides on the other hand were detected 120 and 309 days post-molting in ticks and showed promise for further study. However, the underlying databases used by these investigators to search peptides only included blood species from known sources and thus although they could detect the blood meal source, they did not demonstrate they could taxonomically identify the source.

We were the first research group to apply a protein mass spectrometry-based approach to Chagas disease insect vectors. LC-MS/MS based on hemoglobin protein peptide sequences offers a powerful, quick, and affordable way to evaluate multiple triatomine vector blood meals and does not depend on custom reference libraries, but rather uses publicly available information from GenBank. LC-MS/MS is a powerful tool that can aid in understanding vector ecology and surveillance, frequently outperforming PCR-based methods, leading to valuable data for Ecohealth intervention measures.

1.3 Overview of dissertation

In the subsequent chapters we first demonstrate our ability to identify triatomine vector blood meals from field collected insect vectors in comparison with DNA PCR and 12 S sequencing (manuscript 1, Chapter 2) (Keller et al. 2017). In this proof-of-principle paper we identify hemoglobin peptides from mouse blood samples and four field-collected vectors and identify these to the species level with high accuracy. We verify our blood meal detection with DNA-based 12 S PCR and sequencing and discovered that LC-MS/MS identified more blood meals than the DNA-based method. This was the first application of LC-MS/MS for blood meal identification from field-collected, arthropod disease vectors, and was verified with DNA-based analysis.

In the second paper (manuscript 2, Chapter 3) (Keller et al. 2018), we determine the robustness of our methodology with an experimental feeding study, comparing two proteins – albumin and hemoglobin – against a SINE-PCR-based method. We evaluated our ability to detect blood meals temporally in two experiments – time post-feeding and time post-molting. Hemoglobin-based LC-MS/MS outperformed albumin-based LC-MS/MS and SINE-PCR in both experiments and we were able to detect blood meals to species level up to the maximum time of our experimental samples (4 weeks post-feeding and 12 weeks post-molting). With this data we were able to test the limits of our method experimentally and comparatively.

This then allowed us to scale up our application to a larger number of field-collected insect specimens from Guatemala, Central America, and to identify vector blood meals with a comparison to PCR-based methods (manuscript 3- in preparation,

Chapter 4). We developed an updated, integrative blood meal detection pipeline that allowed for multiple blood meal detection in single vector specimens. Using synthetic AQUA (Absolute QUAntification) peptides as controls, we validated our blood meal species identification pipeline and our ability to detect multiple blood meals. Furthermore, AQUA peptides allowed for the general quantification of a blood meal, showing a positive correlation between the number of peptides identified in a sample and the fmol amount of hemoglobin in a sample. This is the first step in further quantifying blood meals to elucidate species not present in our underlying database. In this study, LC-MS/MS detected almost two-fold the blood meals versus DNA-PCR, and also identified in several samples, additional multiple blood meals in a single vector. Lastly, we show the ability of our methodology to be applied with a lower resolution mass spectrometry instrument, making our blood meal identification technique appealing for limited-resource laboratories that may be more available in some Chagas endemic countries.

Expertise and resources of a research lab, in addition to the goal of a study, factor into choosing appropriate blood meal detection techniques (Kent 2009). Method efficiency is often dependent on the circumstances, while accuracy can be dependent on experimental goals and sample sizes. Ultimately, hemoglobin-based LC-MS/MS vector blood meal identification is a strong and complementary technique to available molecular methods, especially when DNA-based methods fail. Mass spectrometry-based approached are another tool in the toolbox for helping battle neglected tropical diseases and other vector borne diseases.

CHAPTER 2: CHAGAS VECTOR BLOOD MEAL SOURCES IDENTIFIED BY PROTEIN MASS SPECTROMETRY

Judith I. Keller[1], Bryan A. Ballif [1,*], Riley M. St. Clair[1], James J. Vincent[1], M. Carlota Monroy [2,1], and Lori Stevens [1,*]

[1] Department of Biology, University of Vermont, Burlington, Vermont, United States of America

[2] Laboratorio de Entomología Aplicada y Parasitología, Escuela de Biología, Facultad de Ciencias Químicas y Farmacia, Universidad de San Carlos de Guatemala, Edificio T-10 Ciudad Universitaria Zona 12, Ciudad de Guatemala, Guatemala

*Co-corresponding authors:

Email: lori.stevens@uvm.edu (LS), bballif@uvm.edu (BAB)

2.1. Abstract

Chagas disease is a complex vector borne parasitic disease involving blood feeding Triatominae (Hemiptera: Reduviidae) insects, also known as kissing bugs, and the vertebrates they feed on. This disease has tremendous impacts on millions of people and is a global health problem. The etiological agent of Chagas disease, *Trypanosoma cruzi* (Kinetoplastea: Trypanosomatida: Trypanosomatidae), is deposited on the mammalian host in the insect's feces during a blood meal, and enters the host's blood stream through mucous membranes or a break in the skin. Identifying the blood meal sources of triatomine vectors is critical in understanding Chagas disease transmission dynamics, can lead to identification of other vertebrates important in the transmission cycle, and aids management decisions. The latter is particularly important as there is little in the way of effective therapeutics for Chagas disease. Several techniques, mostly DNA-based, are available for blood meal identification. However, further methods are needed, particularly when sample conditions lead to low-quality DNA or to assess the risk of human cross-contamination. We demonstrate a proteomics-based approach, using liquid chromatography tandem mass spectrometry (LC-MS/MS) to identify host-specific hemoglobin peptides for blood meal identification in mouse blood control samples and apply LC-MS/MS for the first time to *Triatoma dimidiata* insect vectors, tracing blood sources to species. In contrast to most proteins, hemoglobin, stabilized by iron, is incredibly stable even being preserved through geologic time. We compared blood stored with and without an anticoagulant and examined field-collected insect specimens stored in suboptimal conditions such as at room temperature for long periods

of time. To our knowledge, this is the first study using LC-MS/MS on field-collected arthropod disease vectors to identify blood meal composition, and where blood meal identification was confirmed with more traditional DNA-based methods. We also demonstrate the potential of synthetic peptide standards to estimate relative amounts of hemoglobin acquired when insects feed on multiple blood sources. These LC-MS/MS methods can contribute to developing Ecohealth control strategies for Chagas disease transmission and can be applied to other arthropod disease vectors.

2.2 Introduction

Vector-borne diseases include some of the most complex disease systems, causing approximately 1.4 million deaths annually worldwide [1]. Chagas disease, a vector-borne neglected tropical disease is endemic in many parts of Latin America. It mostly occurs in communities with limited resources and traditional adobe or baroque houses made out of natural materials [1-3]. Chagas disease claims the lives of an estimated 12,500 people annually, with 8-10 million infected and over 60 million at risk of infection [4-6]. One-third of those infected with the Chagas parasite develop life-threatening illnesses, and it can take up to 20 years to develop diagnosable symptoms, making treatment difficult. Chagas disease is typically transmitted when a blood-feeding triatomine insect vector, also known as a kissing bug, deposits *Trypanosoma cruzi*-laden feces on the skin of a mammalian host. The parasite is subsequently introduced into the blood stream from the insect feces through a break in the skin or through mucous membranes. While congenital transmission occurs in 1-

10% of infants born to Chagas-parasite positive mothers [4], *T. cruzi* transmission to humans occurs primarily during a blood meal from an infected triatomine vector [4, 7, 8].

Identifying blood meal sources of triatomine vectors is critical to understanding Chagas disease transmission dynamics and provides data for evidence-based vector control programs. There is currently no effective vaccine against Chagas disease and although two anti-Trypanosomal drugs, Nifurtimox and Benznidazole, are available, these have considerable side effects and are not always a solution to the overall disease management problem [2, 5, 9]. There are over 140 vector species across the Americas with varying degrees of importance regarding their roles in harboring and transmitting the Chagas parasite to humans [10]. Although the vectors are known to feed on reptiles, birds, and amphibians, the parasite can only reproduce in mammalian hosts, furthering the complexity of interrupting disease transmission [11, 12]. Understanding vector epidemiology and feeding prevalence can be an indicator of how well vector control strategies are working, aiding in Ecohealth control strategies where communities actively participate in reducing the conditions that can increase Chagas transmission [13-15]. Therefore, policies to adjust human behavior and to manage vectors remain more than a partner to medicinal therapeutics, and play increasingly important roles in controlling and preventing infection [14]. Critical to the development of effective policies is an accurate understanding of the sources of vector blood meals.

A complete understanding of the blood meal sources can be challenging for several reasons. DNA-based information is very powerful, relatively inexpensive and

has been used in numerous studies to identify the species of blood meal sources. The relative stability of DNA compared to protein and the ability to PCR-amplify, sub-clone, and sequence DNA provides high sensitivity and specificity in blood meal identification, and therefore these methods are routinely used to determine blood meal sources from insect vectors [11-14, 16-26]. Although some methods can in part address issues such as DNA degradation [27, 28], most DNA-based methods work best with high quality DNA from recently fed vectors [22, 29-31]. Furthermore, given their reliance on an amplification step, DNA-based methods can lead to false positives from contaminating DNA that was not derived from the blood meal itself. Previously employed antibody-based techniques such as the precipitin and antisera tests require fresh material stored at cold temperatures and specific antibodies of possible host species in an area [32-38].

Therefore, additional tools for identification and quantification of insect vector blood meal sources are desirable. Here we demonstrate a liquid chromatography-tandem mass spectrometry (LC-MS/MS) approach based on the identification of highly stable hemoglobin proteins [39-41] which are some of the most abundant proteins in any blood meal [42]. Indeed, the remarkable stability of iron-bound hemoglobin is illustrated by the high amounts of iron and porphyrins derived from hemoglobin using time-of-flight secondary ion mass spectrometry that have been detected in a 46-million-year old fossilized mosquito [39]. In addition, hemoglobin has been detected up to 309 days post-molting under laboratory conditions in ticks through hemoglobin sequence searching [41]. Our approach, based on publicly available DNA and protein sequences,

can be highly precise when using the entirety of published sequences and does not require creating spectral libraries of potential vertebrate hosts as required for other mass spectrometry based techniques that use spectral matching [30]. Because of the high precision, mass spectrometry can be used to accurately identify hemoglobin peptide sequences, many of which are unique to specific vertebrate classes, orders, families, genera, and even species [40-42].

The purpose of this study was to determine if we can apply a proteomic-based, LC-MS/MS approach for identification of blood meal sources of Chagas disease insect vectors. For this purpose, we first used blood from *Mus musculus* (house mouse) to validate our ability to identify mouse hemoglobin peptides from LC-MS/MS to predictions based on a database of hemoglobin sequences we curated through GenBank of the National Center for Biotechnology Information (NCBI) [43]. Subsequently, we provide the first identification of vertebrate host hemoglobin peptides in Chagas disease vectors to determine blood meal sources, using the same LC-MS/MS approach applied to the mouse blood controls. This is the first application of LC-MS/MS for blood meal identification from field-collected, arthropod disease vectors, verified with DNA-based analysis. We confirm our blood meal species identification with DNA analysis based on sequencing the mitochondrial 12S and 16S ribosomal genes and specifically examine insect vectors that have proved challenging for DNA analysis (i.e., stored in ethanol and collected before or after the insect had died). Finally, we demonstrate the potential to estimate absolute quantities of hemoglobin leading the way for further

studies quantifying amounts of hemoglobin from arthropod disease vectors that have fed on multiple taxa.

2.3 Methods

The general workflow to identify blood meal sources is shown in Figure 1. We describe that workflow below.

2.3.1 Ethics Statement

All procedures were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Vermont in accordance with the requirements of the Office of Laboratory Animal Welfare (IACUC protocol 12-045). Mouse blood samples for this study were obtained and treated in accordance with an IACUC-approved protocol encouraging "tissue sharing" of post-mortem tissue.

2.3.2 Collection and storage of mouse blood controls

In order to evaluate LC-MS/MS for blood source identification using hemoglobin peptides, we first validated our ability to identify *M. musculus* blood. Whole blood samples were drawn from three two-month-old Dicer mice [44, 45] immediately following euthanization by cervical dislocation. These transgenic mice with different genetic backgrounds were derived from the cross of C57 and FVB mice [44, 45]. To investigate the role of the anticoagulant ethylenediamine tetraacetic acid (EDTA), blood samples were placed in one of three storage conditions: (1) EDTA (100 mM EDTA; Sigma-Aldrich, Saint Louis, Missouri, USA), (2) dipotassium EDTA (18 mM K_2EDTA; BD Vacutainer, Becton, Dickinson and Company, Franklin Lakes, New

Jersey, USA), and (3) no preservative. Immediately following collection, blood tubes were inverted several times to prevent clotting in the EDTA-containing tubes and then stored at -20 °C for 2 months at which time blood proteins were extracted as described below.

2.3.3 Collection and dissection of insect vectors

Four *Triatoma dimidiata* (Hemiptera: Reduviidae) insect vectors were collected from houses in villages in El Salvador between September and November 2012 (Table 1). Three of these specimens were alive and one was dead at the time of collection (Table 1). Collected insects were stored in 95% ethanol, 5% glycerol, and kept at room temperature for 1-3 months before being stored at -20 °C until dissected in January of 2014. The insect abdomen (consisting of the caudal end of the insect posterior to the crop) was separated into the left and right halves. Specimen halves were randomly assigned to protein or DNA analysis.

2.3.4 Protein extraction, SDS-PAGE, and mass spectrometry

We extracted proteins from *M. musculus* blood by adding 60 µl of mouse blood from the storage conditions listed above, to 20 µl of 95 ˚C 4X denaturing sampling buffer to a final concentration of 5% bromphenol blue, 150 mM Tris pH 6.8, 2% SDS, 5% β-mercaptoethanol, 7.8% glycerol, and heating for 5 min at 95 °C. Each sample was serially diluted with 1X denaturing sampling buffer such that the equivalent of 2 µl of mouse blood was run and analyzed per gel lane. For insect samples, proteins were extracted by adding 100 µl of 95 ˚C denaturing sampling buffer (see above) per 0.1 g of

insect tissue and ground with a clean glass rod in a fume hood. The samples were subsequently heated for 5 min at 95 °C.

Each mouse blood and insect sample was then centrifuged at 16,000 X g for five minutes, and 20 µl of the supernatant was subjected to denaturing 15% SDS-PAGE (37.5 acrylamide: 1 bis-acrylamide) and stained with Coomassie blue. Gel regions surrounding the molecular weight of hemoglobin were excised and diced into 1 mm cubes. In-gel digestion with trypsin, peptide extraction and LC-MS/MS analysis using a linear ion trap-orbitrap (LTQ-Orbitrap; Thermo Electron, Waltham, Massachusetts, USA) was performed as described previously [46] except that all spectra were acquired in the orbitrap for the mouse blood and field-collected insect vector samples.

2.3.5 Using mass spectrometry for the quantification of total hemoglobin and taxon-specific hemoglobin

In some cases, Chagas vectors take blood meals from multiple sources [12]. Therefore, methods that could apportion the contribution of distinct taxa to the total blood meal would be desirable. Successfully accomplishing this using common DNA and mass spectrometry approaches are strongly dependent on well-curated sequence databases. In the case of DNA, taxon-specific PCR primers may be used for targeted amplification analyses to capture qualitative (agarose gel stained PCR products) or quantitative (qPCR) information about the contribution of particular taxa to a blood meal. Additionally, PCR primers specific to regions that are constant across taxa could be used to amplify across a variable region that would then require cloning or Next-

generation sequencing of PCR products in order to attribute the percent any one taxon contributed to the total.

Determining the sequence of amino acids within a peptide by mass spectrometry is predominantly accomplished by matching observed tandem mass spectra with theoretical tandem mass spectra calculated from known protein sequences. While the canonical hemoglobin sequences for most organisms on which Chagas vectors feed are known, it remains a formal possibility that some hemoglobin sequences are not present in current databases; however we based our analysis on hemoglobin because of the stability of the molecule as demonstrated by the ability to identify it over geologic time [39]. Furthermore, it would be desirable to ascertain the contributions of each blood source present in the vector at a given time. A proteomics method to quantify the absolute amount of hemoglobin present in a kissing bug, and the percentage of hemoglobin from any known or unknown species, could take advantage of the fact that some regions of hemoglobins are highly polymorphic while other regions are invariant [41, 47]. The method would involve quantifying the amount of an invariant peptide to determine the total amount of hemoglobin ingested and also quantifying species-specific peptides to ascertain the relative contribution of a given species to the total. If the kissing bug harbored blood from an unknown species then the sum of the species-specific peptides would be less than the amount of the invariant peptide. In proteomics this approach employs stable isotope-containing peptide standards and is known as Absolute QUAntification or AQUA [48].

Quantification using a stable isotope-containing standard peptide was done as described previously [49] using the synthetic AQUA peptide LLVVYPWTQR which contained $^{13}C_5$, $^{15}N_1$-proline synthesized at Cell Signaling Technology (Danvers, Massachusetts, USA). LC-MS/MS methods were as described above, except that only MS1 spectra were acquired in the orbitrap for quantification, while MS2 spectra were collected in the linear ion trap mass spectrometer (LTQ).

2.3.6 Peptide and blood meal identification

LC-MS/MS does not directly sequence the amino acids in a peptide, but gives a spectrum of the masses associated with fragmenting a peptide which can then be matched with high accuracy to expected peptide sequences in an underlying database [50]. Mass spectra were searched using SEQUEST (Thermo Electron V26.12) against a custom forward and reverse concatenated database using a target-decoy approach [51] and allowing for variable oxidation of methionine (+15.9943 Da) and acrylamidation of cysteine (+71.0371 Da). The custom database contained vertebrate hemoglobin sequences (available data for amphibians, reptiles, birds, and mammals) extracted from GenBank on 20 January 2016 with "hemoglobin" in any curated field [43]. We used this strong comparative database with over 17,000 entries to identify hemoglobin peptides in a given sample.

Only doubly- and triply-charged peptide ions were considered. Peptide filtering criteria were: (1) XCorr values greater than or equal to 2.5 (z=2) or 3 (z=3); (2) measured precursor masses +/- 5 PPM; (3) unique ΔCn values greater than or equal to 0.1; and (4) no missed tryptic cleavages except at the extreme N- or C-termini of

peptides where more than one R or K in a row was allowed. These stringent filters resulted in no reverse database matches and thus gave peptide false discovery rates of less than 0.01%. In order to identify the vertebrate species to which the amino acid peptide sequences matched, protein entries from the database were subjected to *in silico* tryptic cleavage using the Pyteomics [52] python tools and library. Taxonomic lineage for each protein was collected from the NCBI taxonomy database [53] (20 January 2016). Peptides were associated with the parent protein taxonomic lineage and stored in a relational database for easy retrieval, allowing us to determine to which vertebrate species an identified hemoglobin peptide matched.

Hemoglobin proteins are sufficiently conserved across these vertebrate taxa to allow mapping of individual peptides to the alpha and beta chains. Peptides are identified with the chain (alpha or beta) and first and last amino acid (i.e. alpha_17-31, or beta_0-8) as mapped to GenBank NP_032244.2, BAG16710.1 for mouse, and P07405.1, XP_003992931.2 and P60529.1, P60524.1 for cat and dog, respectively. However, some variation occurs due to polymorphisms in arginine and lysine residues where the trypsin cuts occur. Among the spectra for a sample, there can be more than one that represent the same peptide. We use the term spectral counts to refer to the number of times a peptides was identified and peptide variant to refer to peptides that map to the same position but vary in a few amino acids. Some hemoglobin peptides are polymorphic within species, others are unique to a particular species, whereas others may be found in multiple species. In addition, some alpha or beta hemoglobin sequences have a variable methionine at the extreme amino-terminal, which

corresponds to alpha_0 or beta_0. If the methionine is absent, the sequence starts with alpha_1 or beta_1 with our designation.

Because some peptides have been reported in multiple species and some polymorphisms may not appear in the curated database, we developed a pipeline (Figure 2) to infer the most likely blood source and verified it with the mouse blood samples. Subsequently we applied the same pipeline for identifying blood meal sources of triatomine insect vectors. The pipeline tabulates the potential taxa represented by the peptides identified by LC-MS/MS in a sample. For identifying the potential blood source of a given sample, we calculated two summary statistics, the first quantifying the protein coverage of a species' hemoglobin by peptide identified in a given sample. The second, based on the spectral count, was the percentage LC-MS/MS spectra of a given sample matching to the hemoglobin peptides identified of a particular species. The taxonomic identification was based on the best match, where the highest matches of the two summary statistics indicate the most likely blood source for each sample (see Figure 2 and Table 2 for more details).

2.3.7 DNA extraction and 12S/16S mitochondrial sequencing of insect abdomens

We verified our results based on LC-MS/MS with a more traditional blood meal identification method for our insect vector samples. From *Triatoma dimidiata* insect abdomen halves, DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) as previously described [19, 54, 55] and identification of blood meal sources was similar to previously published methods based on PCR amplification and sequencing of vertebrate mitochondrial genes [20, 54] except that multiple sets of

primers were used and PCR products were directly sequenced, omitting the cloning step. Briefly, the PCR reaction used primers specific for vertebrate mitochondrial DNA coding for fragments of the 12S and 16S ribosomal RNA gene (hereafter referred to as 12S or 16S primers). Because primer efficiency depends on the blood meal source, three sets of vertebrate 12S and one set of vertebrate 16S primers were used [55-57]. A positive control and negative control (PCR-grade water) were included in each set of PCR amplifications. An ethidium bromide stained, 1.5% agarose gel was used to verify the 215 bp (Kitano), ~100 bp (Melton), 98 bp (Karlsson 16 S), or 111 bp (Karlsson 12 S) PCR products, which were sequenced using BigDye v3.1 (Applied Biosystems, Foster City, CA, USA) and subsequently analyzed with an ABI PRISM 3730xl DNA analyzer (Beckman Coulter, Fullerton, CA, USA). Sequence alignments and editing were done with Sequencher v4.10 (Gene Codes Corporation, Ann Arbor, MI, USA). Taxonomic identification of the sequences was based on the best match of 215 bp (Kitano), ~100 bp (Melton), 98 bp (Karlsson 16 S), or 111 bp (Karlsson 12 S) using the NCBI BLAST algorithm [58].

2.4 Results

In this study, we show that a proteomics-based approach using LC-MS/MS can be a valuable additional technique for identifying blood meal sources using mouse blood to verify the approach as well as identifying blood meals from Chagas disease triatomine insect vectors. First, we show LC-MS/MS is able to unambiguously identify peptides. We then show how these peptide data, combined with our analysis pipeline,

allowed us to correctly identify hemoglobin peptides from *M. musculus* blood to correct animal species with strong support. Finally, we were able to identify dog and cat as blood meal sources from field collected *Triatoma dimidiata* and employ a LC-MS/MS-based method for quantifying a blood meal source.

2.4.1 Mouse Blood

We were able to identify hemoglobin protein peptides unambiguously with our proteomics-based, LC-MS/MS approach (Fig 3, Table 3 and 5). Peptides differing only by a single amino acid, such as YFDSFGDLSSASAIMGNAK and YFDSFGDLSSASAIMGNPK, could be identified with high confidence in our mouse blood (Fig 3). As the mouse blood used in these experiments came from non-isogenic strains of mice with hybrid genetic backgrounds, we could often identify peptide variant for a single peptide position.

Our novel LC-MS/MS method allowed us to correctly identify the taxon of origin when analyzing hemoglobin peptides isolated from *M. musculus* (Table 3). We identified 24 non-redundant hemoglobin protein peptides from the 165 mass spectra collected in the three mouse blood samples. None of the 24 peptides was unique to *M. musculus*, however 22 of the 24 peptides (91.6 %) and 153 of the 165 spectral counts (92.7 %) matched previously reported *M. musculus* hemoglobin sequences in GenBank, making it the most supported blood source species (Fig 2 and S1-S3 Figs). The 22 peptides matching records of *M. musculus* in GenBank matched anywhere from two to as many as 442 taxa. Of the two peptides (0-2 per sample) where the top match was not *M. musculus*, one occurred at the alpha_17-31 position and the other at beta_18-30.

However, for these two peptides, the number of amino acids matching *M. musculus* as the blood source was high (Table 4).

In terms of protein coverage, we found 16 of the 20 (80 %) expected tryptic peptides (greater than 4 amino acids, Fig 4). Trypsin digestion is expected to yield 9 alpha and 11 beta hemoglobin peptides greater than 4 amino acids in length from *M. musculus* hemoglobin based on GenBank entries NP_032244 and BAG16710. Of the 16 peptides that we identified, one (alpha_17-31) had two peptide variants (Table 3). When blood was stored in EDTA (samples 1996, 2308), more non-redundant peptides were identified, while spectral counts were similar in all three treatments (Table 3).

2.4.2 Identification of blood meal sources of triatomine insect vectors

Our LC-MS/MS method identified blood meals from insects collected alive and dead; however, blood meals were not detected with DNA sequencing methods for three of the four samples. By LC-MS/MS, Canidae species, especially dog, was the most strongly supported blood meal source in three (076, 101, 112) of the field collected samples and cat in the other (051) (Table 4 and S4-S7 Figs). These results are based on 22 non-redundant hemoglobin peptides although the number of non-redundant peptides per samples was 50% fewer (8 for the blood meal from the insect compared to 12 for mouse blood) and the number of spectra per sample was at most half of the mouse blood (3-22 for the blood meal from the insect compared to 54-56 for mouse blood) (Table 5). Two of the 7 peptides from sample 051 were unique to cat, and 6 of the 7 (85.7%) matched cat hemoglobin sequences in GenBank (S4 Fig). The next closest match was to two members in the squirrel (Sciuridae) family (14 %). None of

remaining peptides from the other three samples were unique to dog, and all except one were previously reported in dog (86.2 – 100% peptides matching, Table 4). Spectral count matching to the top identified species was also high across the four samples (92.3 – 100%, Table 4). The next closest matches were *Canis latrans,* and *Chrysocyon. brachyurus,* both members of the Canidae family. In terms of amino acids identified, >99% of amino acids identified matched the most likely blood source(s) (Table 4).

Although the match to previously reported hemoglobin sequences in GenBank was high, the number of non-redundant peptides and the peptide coverage was low (range 3-14 peptides with 13.6 % - 28 % coverage, Fig 5), demonstrating that only a few peptides are sufficient to identify a blood meal. Note that for cat there are 10 alpha and 12 beta predicted hemoglobin tryptic peptides and for dog there are 9 alpha and 13 beta predicted tryptic peptides, considering peptides greater than 4 amino acids. Of the 22 peptides identified in triatomine insect vectors, only two peptides were variable within species, alpha_62-90 in sample 101, and beta_66-76 in sample 051. Of the insect vectors collected alive (051, 101, 112), more non-redundant peptides as well as spectral counts were identified (Table 5). Taxonomic resolution was lower for the sample collected dead, only 3 peptides were identified. Still, the blood meal source could be resolved to three possible species (*C. lupus familiaris, C. latrans, C. brachyurus*) in the Canidae family.

While LC-MS/MS was able to identify blood meals in all four insects, blood meals were not identified with more traditional 12S/16S mitochondrial DNA sequencing methods for three of the four samples because there was no visible band

from the PCR reaction, while positive controls amplified successfully. Our previous studies demonstrated lack of PCR product indicates no recent blood meal rather than PCR inhibition [13, 22] and experimental studies have revealed that DNA detection can drop off as early as 1-2 weeks after a blood meal depending on blood meal host species [59]. The DNA sequence and LC-MS/MS results did agree for the one insect where they were both available (Table 6). Sample FER112-2 had a visible PCR product for 3 primer sets (the 12S Karlsson primers did not amplify), and matched *C. lupus familiaris* (domestic dog) 100%, 100% and 99% for 12S Kitano, 16S Karlsson and 12S Melton primers, respectively.

2.4.3 Quantification of blood meal sources of triatomine insect vectors

As a first-step in developing a method for quantifying blood meals, we demonstrated the quantification of a highly-conserved mammalian hemoglobin peptide arising from the blood meal of a field-collected kissing bug. To tryptic peptides derived from the insect blood meal was added a highly conserved synthetic tryptic peptide from beta-hemoglobin, LLVVYPWTQR, which harbored a $^{13}C_5$, $^{15}N_1$-labeled proline residue ($\Delta M = 6.012$). As expected, both the insect blood meal-derived peptide and the synthetic standard eluted at the same time enabling a direct comparison of the relative abundance of each in precursor (MS1) scans. Furthermore, the MS2 scans of the native blood meal and synthetic peptides unambiguously distinguish the two peptides as their spectra differ only in peptide fragments containing the heavy-labeled peptide (Fig 6). In this example, the relative abundance of the native peptide is approximately 50% of the

synthetic peptide; the heavy-to-light ratio of synthetic to native peptide was 1.94:1 ($4.55x10^7$:$2.35x10^7$ comparing monoisotopic peak intensities).

2.5 Discussion

We demonstrate the application of LC-MS/MS for the identification of blood sources based on hemoglobin peptides. After validating our methodology with mouse blood control samples, we identified blood meal sources of Chagas disease insect vectors. This is the first application of LC-MS/MS to field-collected arthropod disease vectors, the first use of LC-MS/MS in this disease system, and the first comparison of LC-MS/MS to DNA-based methods. We were able to detect blood meals in three insect vectors where nothing was detected using DNA-based methods and we showed the potential to quantify a blood meal using synthetic AQUA peptides. These results are important because understanding blood meal composition can lead to the development of Ecohealth control strategies for Chagas disease transmission, aiding in understanding and managing disease transmission dynamics.

2.5.1 Verification with mouse blood

Our findings show LC-MS/MS to be a valuable tool in addition to DNA analysis for identification of blood meal sources in triatomine disease vectors. We used mouse blood to verify our ability to detect hemoglobin protein peptides with LC-MS/MS. We found the highest match to *M. musculus*; over 90 % of the peptide positions along alpha and beta hemoglobin and over 90 % of the spectra, matched *M. musculus*-reported sequences in GenBank. There is evidence that preservation methods

influenced the number of peptides recovered from a blood sample. Twice as many peptide variants and more non-redundant peptides were recovered from both samples stored in EDTA.

Across the mouse blood samples, the only two peptides that did not match *M. musculus* hemoglobin entries currently in GenBank were alpha_17-31 and beta_18-30. All three control mouse blood samples identified both the expected version of alpha 17-31 (IGGHGAEYGAEALER) matching *M. musculus*, while one blood samples (2308) had an unexpected variant (VGSHAGEYGAEALER) currently reported in two species in the Ochotonidae (pika) family, *Tapirus terrestris* (Brazilian tapir), and *Microtus oeconomus* (root vole). An undescribed polymorphism in *M. musculus* is the most probable explanation for this non-match.

The second non-match peptide, beta_18-31 peptide NVADEVGGEALGR, matched another rodent, *Tamiasciurus hudsonicus*, the American red squirrel, and occurred in two mouse blood samples (2308, 2310). Although not detected with our current stringent filtering, the very similar peptide VNADEVGGEALGR differing by the inversion of the first two amino acids on the N-terminal end does match *M. musculus*. The detection of beta_18-31 depends on the presence and correct identification of b1 (N or V) and b2 (NV or VN) ions (and corresponding y ions) in the MS2 spectra. Generally, the lowest *m/z* b-type and highest *m/z* y-type ions are of lower abundance which could have led to the misidentification of this peptide. NVADEVGGEALGR (or VNADEVGGEALGR) is routinely observed with a relatively high signal. However, determination of the sequence at the extreme N-

terminus is challenging. Since the blood source in this case is known, *M. musculus*, we infer that this peptide is indeed from mouse.

In spite of potential unreported polymorphisms or potential misattribution by the mass spectrometry search software, our results still support the top blood source match with high likelihood. When quantifying the percent match of the amino acid sequences of the peptides identified, 98 % (2310) and 96 % (sample 2308) of amino acids identified support *M. musculus* as the correct blood source. This is comparable to the DNA best match sequence identification (99-100%).

2.5.2 Blood meal identification from field-collected insect vectors

With the *T. dimidiata* insect samples, we provide the first analysis of field-collected specimens with LC-MS/MS. Similar to the analysis of the mouse control samples, almost all of the peptides identified from the field-collected specimens matched a single blood source. Of the 22 non-redundant peptides identified from 50 spectral counts, only two peptides (2/50 spectral counts, 4 %) had not been previously reported in the taxon with the strongest support. For sample FER051, beta_66-76 had previously been reported in two species in the squirrel (Sciuridae) family and differed by one amino acid from the fully tryptic peptides previously reported for cat, (VLDSFSDGLK in squirrel, VLNSFSDGLK in cat, Table 5). In addition to a polymorphism, this sequence could arise from a chemical conversion such as deamidation [60]. Overall, the percentage of amino acids identified that matched the top species, *F. catus* was 99.5 %. For sample FER101, one peptide variant of alpha_62-90 was previously reported in *Leptonychotes weddellii* (Weddell seal) and also differed

by only one amino acid from the peptide matching dog, the most highly supported blood source for that particular sample (VADALTTAVAHLDDLPGALSALSDLHAYK for dog, VADALTTAVSHIDDLPGALSALSDLHAYK for seal, Table 5). Of note, standard mass spectrometry analysis does not differentiate between leucine (L) and isoleucine (I) as these two amino acids have the same mass to charge ratio [61].

The amino acid difference for sample FER101 is in alpha hemoglobin site 71, one of the most variable hemoglobin amino acid positions in *Peromyscus* species [47]. Because there was only one protein sequence for alpha and beta hemoglobin from a domestic dog in GenBank when we made our database, an unreported variation in dog is the likely explanation for the difference. Polymorphisms in hemoglobin are known to expand along an altitudinal gradient [41, 47, 62] and are a more likely explanation than the insect vector having fed on seal at least 10,000 kilometers away in Antarctica. The villages in El Salvador where these samples were taken are in more rural areas, so a previously unknown polymorphism in the local dog population is a possibility. Again, 99.5 % of the amino acid sequence of the peptides identified matched the top match blood meal source, *C. lupus*, supporting the high likelihood of the triatomine indeed having fed on domestic dog.

In spite of the fact that triatomine vectors digest blood meals slowly (up to 10 weeks) [14, 16, 24, 29], DNA-based methods for identifying blood meal sources fail in upwards of 50 % of Chagas insect vectors [23, 24, 33-35, 37, 38, 63, 64]. In this regard, hemoglobin is a particularly stable protein and appears to last longer than DNA

signatures [39]. Hemoglobin protein-based LC-MS/MS methods could be used when a sample does not show results with traditional DNA methods.

Interestingly, three of the four vectors took blood meals from dogs. Domestic dogs are considered the most important domestic reservoir for Chagas disease, although they may serve as a bridge species as well, connecting sylvatic and domestic cycles [65-67]. Previous studies have shown that *T. infestans* vectors are more likely to take a blood meal from a dog than a human [2, 68, 69], however keeping approximately two dogs in the household can increase human disease prevalence [2]. The results of our study of *T. dimidiata* from El Salvador supports previous empirical data that triatomine insect vectors feed on domestic dogs [65, 68-72]. One of the four vectors collected in the domestic environment within the house fed on cat. Cats as blood meal hosts are thought to occur much less frequently than dogs [2, 19, 73]. Still, cats can serve as both a blood meal for the vectors and host for *T. cruzi* [66, 68].

Finally, we demonstrate the potential of mass spectrometry-based methods to quantitatively profile triatomine blood meal sources. Using a stable isotope-labeled synthetic standard we show the quantification of a specific hemoglobin peptide derived from a field-collected sample (Fig 6). Using defined amounts of synthetic standards enables the absolute quantification of a given peptide in a sample. This could be a useful approach to quantify both absolute amounts of hemoglobin generally, as well as relative hemoglobin amounts from specific taxa when a vector has acquired blood meals from more than one species. For example, if a triatomine fed on both a dog and a human, there exist hemoglobin peptides that are common to the two mammals, as well

as peptides that are unique to each. If 100 fmol of total hemoglobin was detected by measuring peptides in common between the two species, but species-specific peptides measured 30 fmol of dog peptides to 70 fmol of human peptide one could estimate the relative contribution of each to the blood meal derived from the vector. Additionally, if the sum of the two species-specific peptides fell short of the 100 fmol, conservation of mass would require the presence of a third blood meal, or at least an unanticipated polymorphism. The ability to quantify species-specific blood meals would help determine relative feeding prevalence of triatomines in an area with multiple potential hosts, and by extension would help focus management strategies based on particular host species.

Although LC-MS/MS presents a valuable tool in addition to DNA-based techniques, its cost could limit its application to situations where DNA analysis fails on a significant number of samples [74]. The upfront cost of LC-MS/MS platforms can be substantial, however the analysis itself can be relatively inexpensive given that these platforms typically exist in a multi-user facility with institutional support, as described previously [30] (S3 Table). In addition, LC-MS/MS can be accessible to researchers anywhere with such platforms being widely distributed, similarly to DNA sequencing facilities [30]. Costs per sample for a single LC-MS/MS run can range anywhere from $10 to $100 per sample or more when using multi-user facilities with proteomic platforms (S3 Table). Successful application of LC-MS/MS depends on a large and robust underlying database of hemoglobin protein sequences, just as DNA analysis depends on available DNA sequences. It is possible that a species would not be

represented in the database and therefore not identified. However, the high conservation of hemoglobin would likely permit the identification of any blood meal to at least an evolutionarily-related species [30].

Our study detailing a protein-based mass spectrometry approach demonstrates proof-of principle support for identifying blood meal sources in Chagas disease insect vectors using hemoglobin peptides with surprisingly species-level resolution. Knowing how blood meals are changing in endemic areas is critical for vector control. Information such as blood meal prevalence leading to adjusting policies to manage potential animal hosts and insect vectors play increasingly important roles in controlling disease and preventing parasite infection, as vector control is still one of the most efficient control strategies [2, 14]. The experiments presented here show that using LC-MS/MS is a sensitive, valuable methodology to identify blood meal sources in Chagas disease triatomine insect vectors.

2.6 Acknowledgements

We thank the El Salvador Ministry of Health for insect sample collection and Kristiaan Finstad and the UVM Stein Lab for providing mouse blood control samples. Thank you to Jeffrey Knott and Jason Reynolds at Cell Signaling Technology for synthesizing the stable isotope-containing peptides. Thanks to Marion Weir for helpful guidance with mass spectrometry analyses.

2.7 Figures and Tables

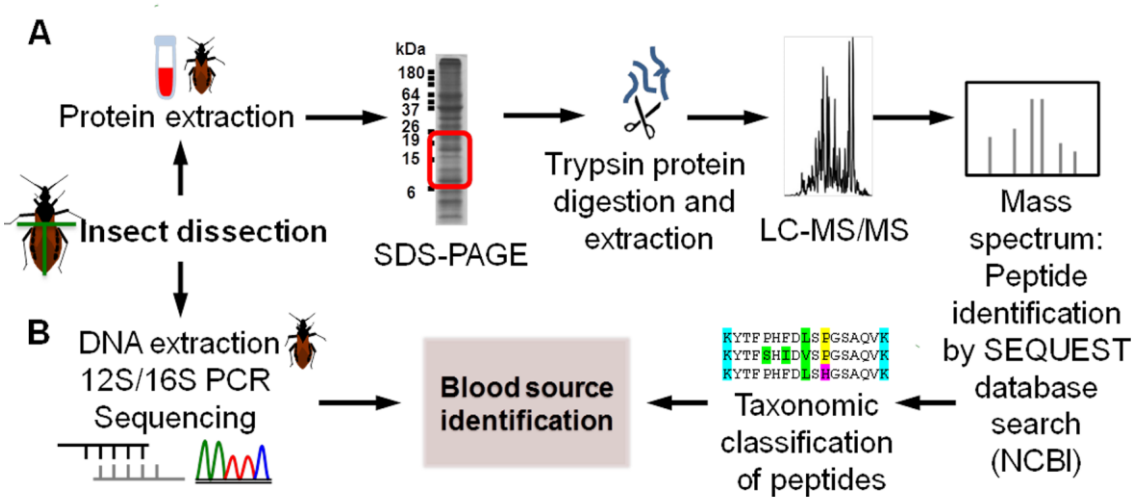

Figure 2.1: General workflow describing LC-MS/MS-based and DNA-based identification of blood meal sources. Insects were dissected into left and right abdomen, and hemoglobin peptides from (A) mouse blood as well as triatomine insect vectors were identified with LC-MS/MS, while (B) triatomine insect vector blood meals were also identified with DNA-based methods for comparison.

Table 2.1: *T. dimidiata* insect vectors collected from houses in villages in El Salvador, abdomen halves assigned to LC-MS/MS protein or DNA analysis.

Sample	Analysis	Sex	Collection Location			Collection
			Village	Subvillage	Ecotope	
FER 0051-1	LC-MS/MS	M	Azacualpa	El Tablón	domestic	alive
FER 0051-2	DNA					
FER 0076-1	DNA	M	Azacualpa	Ocotillo	domestic	dead
FER 0076-2	LC-MS/MS					
FER 0101-1	DNA	M	Cañaverales	Dolores	peridomestic	alive
FER 0101-2	LC-MS/MS					
FER 0112-1	LC-MS/MS	F	Cañaverales	Plantanares	peridomestic	alive
FER 0112-2	DNA					

Sample: 1996 non-redundant peptides identified in sample

A

M. musculus
NP_032244.2, BAG16710.1

	alpha_17-31	alpha_32-40	alpha_41-56	alpha_91-99	alpha_128-139	beta_1-8	beta_18-30	beta_41-59		beta_66-82	beta_83-95	beta_105-120		beta_121-132	beta_133-144	beta_133-146	**Total**
no. amino acids/peptide	15	9	16	9	12	8	13	16	16	17	13	16	16	12	12	14	214
no. peptide variants	1	1	1	1	1	1	1	2		1	1	2		1	1	1	16
spectral count	1	2	2	1	1	1	5	26	5	1	1	2	1	3	1	1	54

B taxanomic affiliations

																	range
no. of classes	1	1	1	3	1	1	1	1	1	1	1	1	1	1	1	1	(1- 3)
no. of orders	1	1	2	50	8	2	1	1	1	1	3	1	1	2	7	6	(1 - 51)
no. of families	6	2	7	127	27	4	1	1	1	1	6	1	1	4	16	15	(1 - 128)
no. of genera	7	2	11	290	62	8	1	1	2	1	14	2	1	5	26	25	(1 - 291)
no. of species	9	2	19	442	93	12	3	8	7	4	32	8	3	13	44	43	(2 - 443)

Species reported with peptide																
Mus musculus	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Mus spicilegus							x	x	x	x	x	x	x	x	x	x
Mus macedonicus							x	x	x		x	x	x	x	x	x
Mus spretus								x		x	x	x		x		
Cricetomys gambianus	x	x	x	x	x											
Mus pahari										x					x	x
no. species not listed	5		15	434	88	9		4	4		28	4		7	38	37

Figure 2.2: Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for mouse blood sample 1996. (A) Peptides identified by SEQUEST mapped to alpha and beta hemoglobin, using GenBank NP_032244.2, BAG16710.1 reference alignments for *M. musculus* hemoglobin (amino acid 0-141 for alpha and 0-146 for beta hemoglobin). Peptides are named based on hemoglobin chain and the position of the first and last amino acid (i.e. alpha_17-31). When more than one peptide mapped to a position (e.g., beta_41-59), corresponding data for both peptide variants are shown. (B) Taxonomic affiliations for each peptide that matched five species or less are tabulated, in this case peptide alpha_32-40 matched two species (orange highlight), two peptides matched three species (yellow), and one peptides matched four species (blue). The peptides are then examined to determine if they are consistent with each possible species. 'x' represents a peptide that matched the particular species in question.

Table 2.2: Using the information from the pipeline (Figure 2), two summary statistics quantify the strength of blood meal source identification. The percentage of peptides identified quantify whether a particular peptide does/does not match the proposed blood source. The percent of spectral counts matching a particular taxon provides a semi-quantitative measure of relative abundance. The species with the highest summary statistics is determined to be the most likely blood source, in this case *M. musculus* (green highlight).

Species reported with peptide	Total peptide matches per taxon	Total peptide non-matches per taxon	Percent peptides identified matching	Percent spectral count matching
Mus musculus	16	0	100.0%	100.0%
Mus spicilegus	10	6	62.5%	85.2%
Mus macedonicus	9	7	56.3%	83.3%
Mus spretus	5	11	31.3%	61.1%
Cricetomys gambianus	5	11	31.3%	13.0%
Mus pahari	3	13	18.8%	5.6%

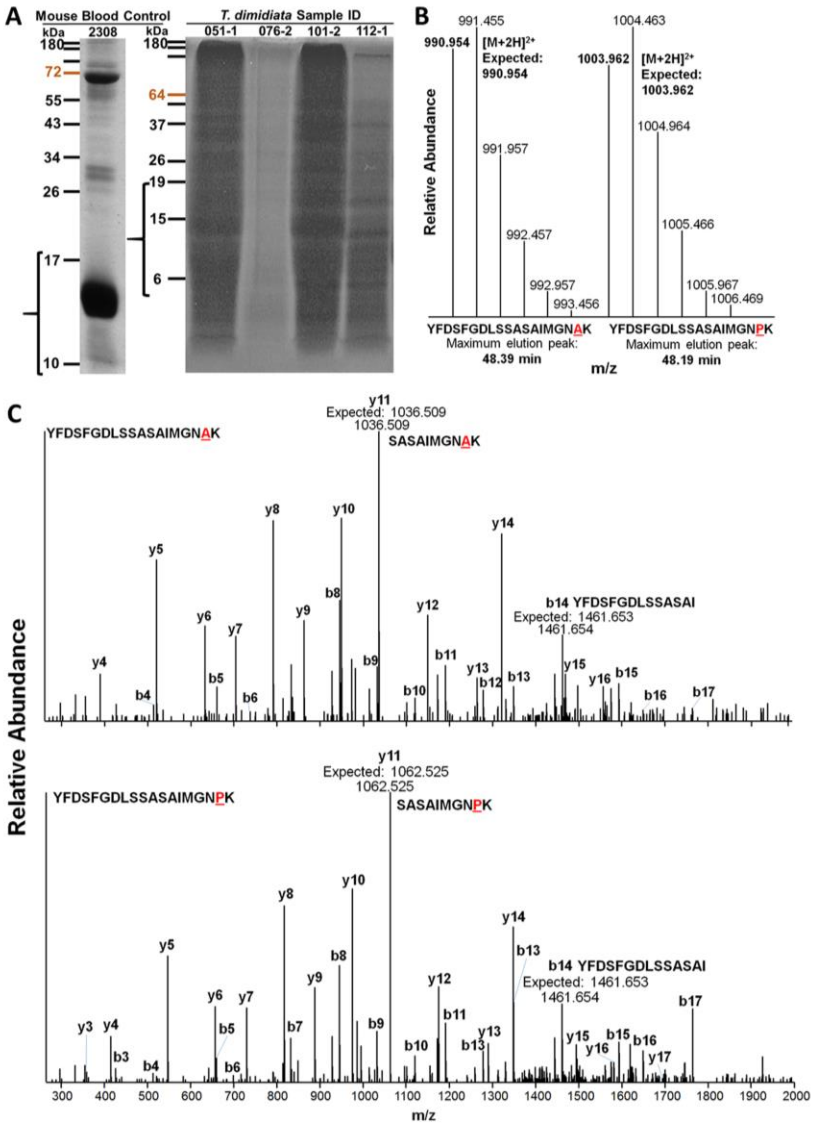

Figure 2.3: Proteomics-based, LC-MS/MS distinguishes unique but nearly identical peptides. (A) Representative SDS-PAGE results of mouse blood samples and of four *T. dimidiata* insect vectors. (B) MS1 spectra of two doubly-charged peptide ions differing by a single amino acid- YFDSFGDLSSASAIMGNAK and YFDSFGDLSSASAIMGNPK, identified in mouse blood control sample 2308. Mass to charge ratios for these peptides differ only by the difference between the variable amino acid (alanine or proline). The two peptides also elute at slightly different times. (C) Low energy collision-induced dissociation fragmentation (MS2) mass spectra of the aforementioned peptide ions allow determination of the peptide sequence. Peaks are labelled as per convention with b-type fragment ions (those derived from the amino terminus) and y-type fragment ions derived from the carboxyl-terminus. Given the variable amino acid is the penultimate carboxyl-terminal amino acid, y2 and higher y-type ions differ by the mass variability between alanine and proline (e.g. y11), while almost all b-type ions (e.g. b14) show equal *m/z* measurements. Expected and observed masses for identified fragment ions can be found in Supplementary Tables S8/9.

Table 2.3: Tryptic peptides identified in mouse blood control samples, including hemoglobin position and spectral counts. [a] indicates peptides not matching the known blood sources, *M. musculus*; 1996, 100 mM EDTA; 2308, 18 mM K_2EDTA; 2310, no preservative

Peptide	Hemoglobin position	Mouse blood control sample		
		1996	2308	2310
		Spectral Counts		
IGGHGAEYGAEALER	alpha_17-31	1	2	1
VGSHAGEYGAEALER [a]	alpha_17-31		1	
MFASFPTTK	alpha_32-40	2	1	
TYFPHFDVSHGSAQVK	alpha_41-56	2	1	
VADALANAAGHLDDLPGALSALSDLHAHK	alpha_62-90			5
LRVDPVNFK	alpha_91-99	1		
FLASVSTVLTSK	alpha_128-139	1		
VHLTDAEK	beta_1-8	1		
AAVSGLWGK	beta_9-17			6
NVADEVGGEALGR [a]	beta_18-30		11	11
VNSDEVGGEALGR	beta_18-30	5	3	
YFDSFGDLSSASAIMGNAK	beta_41-59	26	24	27
YFDSFGDLSSASAIMGNPK	beta_41-59	5	3	
KVITAFNDGLNHLDSLK	beta_66-82	1	2	
VITAFNEGLK	beta_67-76		1	
VITAFNDGLNHLDSLK	beta_67-82		5	
GTFASLSELHCDK	beta_83-95	1		
LHVDPENFR	beta_96-104		1	
LLGNAIVIVLGHHLGK	beta_105-120	1		
LLGNMIVIVLGHHLGK	beta_105-120	2	1	
DFTPAAQAAFQK	beta_121-132	3		1
VVAGVAAALAHK	beta_133-144			4
VVAGVATALAHK	beta_133-144	1		
VVAGVATALAHKYH	beta_133-146	1		
No. of non-redundant peptides		16	13	7
No. of peptide positions	16	13	8	7
No. of spectral counts		54	56	55

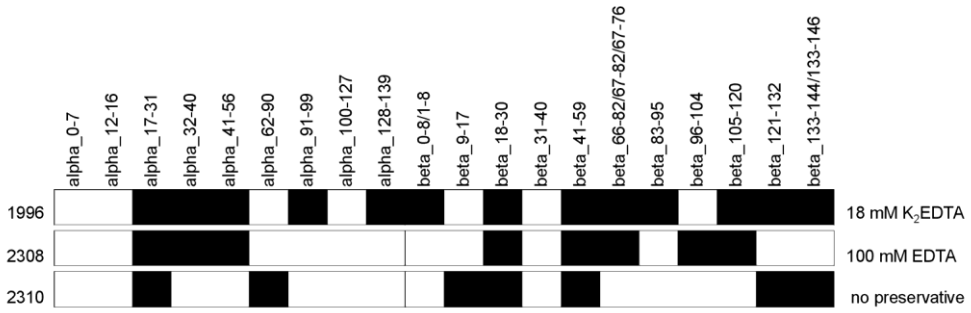

Figure 2.4: Hemoglobin protein peptide coverage from identified tryptic peptides for three mouse blood samples. Three mouse blood samples stored in varying amounts of EDTA are shown. The peptides are shown with equal width, not in proportion to the length of the peptide, and only peptides greater than four amino acids are shown.

Table 2.4: Summary statistics and blood source identifications of mouse blood controls and triatomine insect samples. 1996, 100 mM EDTA; 2308, 18 mM K2EDTA; 2310, no preservative

Sample	Blood source identification	Percent peptides identified matching	Percent spectral count matching	Percent amino acids identified matching
Mouse blood samples				
1996	*M. musculus*	100.0%	100.0%	100.0%
2308	*M. musculus*	84.6%	78.57%	96%
2310	*M. musculus*	85.7%	80.0%	98%
***T. dimidiata* insect samples**				
051	*F. catus*	85.7%	92.3%	99.5%
076	*C. lupus familiaris*, *C. latrans*, or *C. brachyurus*	100.0%	100.0%	100.0%
101	*C. lupus*	92.3%	95.7%	99.5%
112	*C. lupus*	100.0%	100.0%	100.0%

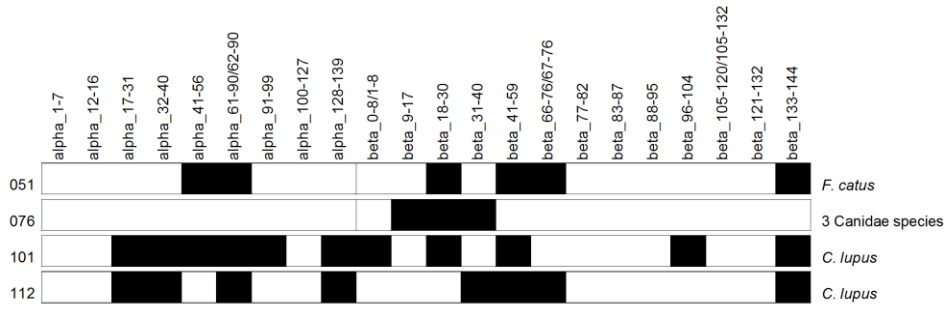

Figure 2.5: Hemoglobin protein peptide coverage from identified tryptic peptides from blood meals of four triatomine insects. Triatomine insect vectors were collected alive (051, 101, 112) and dead (076). The peptides are shown with equal width, not in proportion to the length of the peptide, and only peptides greater than four amino acids are shown.

Table 2.5: Peptides identified in triatomine insects, including hemoglobin position and spectral counts. [a] indicates peptides not matching the most likely blood meal (051, cat; 076, 101, 112, dog). [b] indicates a peptide that was added to GenBank 29 Dec 2016, after the original database was created.

Peptide	**Hemoglobin position**	***T. dimidiata* samples**			
		FER051	FER076	FER101	FER112
		Spectral Counts			
IGGHAGDYGGEALDR	alpha_17-31			1	1
TFQSFPTTK	alpha_32-40			1	2
TYFPHFDLSHGSAQVK	alpha_41-56	1			
TYFPHFDLSPGSAQVK	alpha_41-56			3	
KVADALTTAVAHLDDLPGALSALSDLHAYK	alpha_61-90			1	
VADALTQAVAHMDDLPTAMSALSDLHAYK	alpha_62-90	4			
VADALTTAVAHLDDLPGALSALSDLHAYK	alpha_62-90			2	2
VADALTTAVSHIDDLPGALSALSDLHAYK [a]	alpha_62-90			1	
LRVDPVNFK	alpha_91-99			1	
FFAAVSTVLTSK	alpha_128-139			3	1
VHLTAEEK	beta_1-8			3	
SLVSGLWGK	beta_9-17		1		
VNVDEVGGEALGR	beta_18-30	1	1	1	
LLIVYPWTQR	beta_31-40		1		1
FFDSFGDLSTPDAVMSNAK	beta_41-59			3	3
FFQSFGDLSSADAIMSNSK	beta_41-59	3			
KVLDSFSDGLK [a]	beta_66-76	1			
VLNSFSDGLK	beta_67-76	2			1
LHVDPENFK	beta_96-104			1	
VVAGVANALAHK	beta_133-144			1	1
VVAGVASALAHR [b]	beta_133-144	1			
No. of non-redundant peptides		7	3	14	8
No. of peptide positions	16	6	3	12	8
No. of spectral counts		13	3	22	12

Table 2.6: Comparison of LC-MS/MS-based and DNA-based blood meal source identification methods.

Sample	12S/16S blood meal identification	LC-MS/MS blood meal identification
FER 051	None	*F. catus*
FER 076	None	3 Canidae species
FER 101	None	*C. lupus familiaris*
FER 112	*C. lupus familiaris*	*C. lupus familiaris*

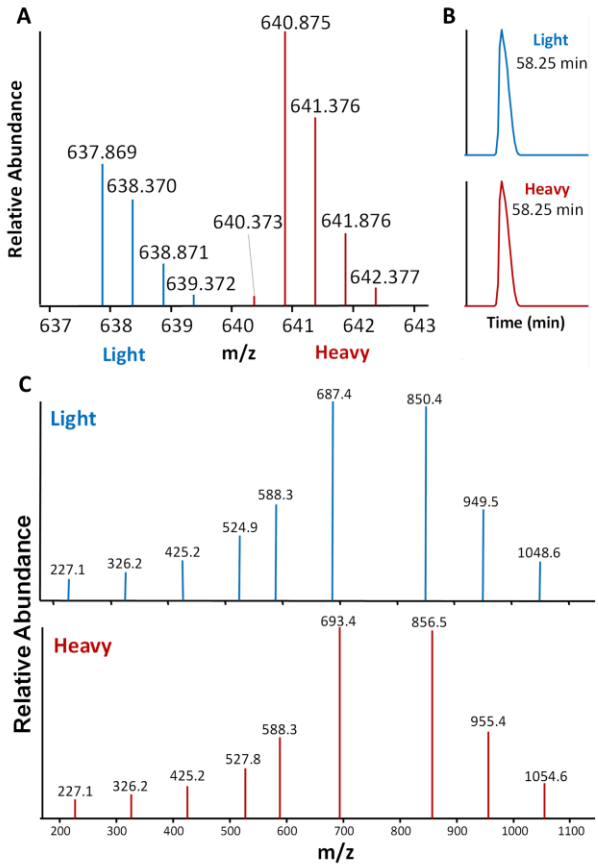

Figure 2.6: Stable isotope-labeled standard facilitates quantification of native hemoglobin peptide. (A) MS1 spectrum showing the relative abundance of a native hemoglobin peptide derived from the blood meal of a field-collected kissing bug (blue isotopic envelope) in comparison to an introduced synthetic peptide standard (red isotopic envelope) of the same sequence containing a $^{13}C_5$, $^{15}N_1$-labeled proline. (B) Partial chromatogram showing the co-elution of the monoisotopic masses of both the native "light" and synthetic "heavy" peptide ions shown in A. (C) Low-energy collision-induced dissociation MS2 fragmentation spectra of the doubly-charged precursor ions for the native and synthetic peptides.

S1 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for mouse blood 1996.

S2 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for mouse blood 2308.

S3 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for mouse blood 2310.

S4 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for *T. dimidiata* insect vector sample 051.

S5 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for *T. dimidiata* insect vector sample 076.

S6 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for T. dimidiata insect vector sample 101.

S7 Fig. Pipeline for inferring blood meal sources from hemoglobin peptides illustrated for *T. dimidiata* insect vector sample 112.

S1 Table. Ion table providing the expected and observed *m/z* values for singly-charged fragment ions for peptide ion YFDSFGDLSSASAIMGNPK derived from spectrum shown in Fig 3C. Proline is underlined as it is the variable amino acid between the two peptides ions described in Fig. 3C.

S2 Table. Ion table providing the expected and observed m/z values for singly-charged fragment ions for peptide ion YFDSFGDLSSASAIMGNAK derived from the spectrum shown in Fig 3C. Alanine is underlined as it is the variable amino acid between the two peptides ions described in Fig. 3C.

S3 Table. Per sample cost of LC-MS/MS platform in four proteomics core facilities. Prices for the University of Vermont Proteomics Core Facility are for the high-resolution linear ion trap-orbitrap (LTQ-Orbitrap; Thermo Electron, Waltham, Massachusetts, USA) used for this study. Prices shown for other facilities are for comparable instruments. Self-run LC-MS/MS platforms can be run for as low as $4.75 per sample (see Önder et al. 2013, Supplementary Table S3 for further details).

2.8 References

1. Campbell-Lendrum D, Manga L, Bagayoko M, Sommerfeld J. Climate change and vector-borne diseases: what are the implications for public health research and policy? Philosophical transactions of the Royal Society of LondonSeries B, Biological sciences. 2015;370(1665):10.1098/rstb.2013.0552. doi: 10.1098/rstb.2013.0552 [doi].
2. Cohen JE, Gürtler RE. Modeling household transmission of American trypanosomiasis. Science (New York, NY). 2001;293(5530):694-8. doi: 10.1126/science.1060638 [doi].
3. Zeledon R, Rabinovich JE. Chagas disease: an ecological appraisal with special emphasis on its insect vectors. Annual Review of Entomology. 1981;26(1):101-33.
4. Bern C, Kjos S, Yabsley MJ, Montgomery SP. Trypanosoma cruzi and Chagas' Disease in the United States. Clinical Microbiology Reviews. 2011;24(4):655-81. doi: 10.1128/CMR.00005-11 [doi].
5. Dumonteil E, Herrera C. Ten years of Chagas disease research: Looking back to achievements, looking ahead to challenges. PLoS Neglected Tropical Diseases. 2017;11(4):e0005422.
6. Rassi A, Marin-Neto JA. Chagas disease. The Lancet. 2010;375(9723):1388-402.
7. Hotez PJ, Dumonteil E, Cravioto MB, Bottazzi ME, Tapia-Conyer R, Meymandi S, et al. An unfolding tragedy of Chagas disease in North America. PLoS Negl Trop Dis. 2013;7(10):e2300.
8. Sánchez-Valdéz FJ, Pérez Brandán C, Ferreira A, Basombrío MÁ. Gene-deleted live-attenuated Trypanosoma cruzi parasites as vaccines to protect against Chagas disease. Expert Review of Vaccines. 2015;14(5):681-97.
9. Coura JR, De Castro SL. A critical review on Chagas disease chemotherapy. Memórias do Instituto Oswaldo Cruz. 2002;97(1):3-24.
10. Justi SA, Galvão C. The Evolutionary Origin of Diversity in Chagas Disease Vectors. Trends in Parasitology. 2017;33(1):42-52. doi: https://doi.org/10.1016/j.pt.2016.11.002.
11. Lucero DE, Ribera W, Pizarro JC, Plaza C, Gordon LW, Pena R, Jr., et al. Sources of Blood Meals of Sylvatic Triatoma guasayana near Zurima, Bolivia, Assayed with qPCR and 12S Cloning. PLoS Negl Trop Dis. 2014;8(12):e3365. doi: 10.1371/journal.pntd.0003365. PubMed PMID: 25474154; PubMed Central PMCID: PMC4256209.
12. Waleckx E, Suarez J, Richards B, Dorn PL. Triatoma sanguisuga blood meals and potential for Chagas disease, Louisiana, USA. Emerg Infect Dis. 2014;20:2141-3.
13. Pellecer MJ, Dorn PL, Bustamante DM, Rodas A, Monroy MC. Vector blood meals are an early indicator of the effectiveness of the Ecohealth approach in halting Chagas transmission in Guatemala. The American Journal of Tropical Medicine and Hygiene. 2013;88(4):638-44. doi: 10.4269/ajtmh.12-0458 [doi].

14. Lucero DE, Morrissey LA, Rizzo DM, Rodas A, Garnica R, Stevens L, et al. Ecohealth Interventions Limit Triatomine Reinfestation following Insecticide Spraying in La Brea, Guatemala. The American Journal of Tropical Medicine and Hygiene. 2013;88(4):630-7. doi: doi:https://doi.org/10.4269/ajtmh.12-0448.
15. Boischio A, Sánchez A, Orosz Z, Charron D. Health and sustainable development: challenges and opportunities of ecosystem approaches in the prevention and control of dengue and Chagas disease. Cadernos de Saúde Pública. 2009;25:S149-S54.
16. Mota J, Chacon JC, Gutiérrez-Cabrera AE, Sanchez-Cordero V, Wirtz RA, Ordonez R, et al. Identification of blood meal source and infection with Trypanosoma cruzi of Chagas disease vectors using a multiplex cytochrome b polymerase chain reaction assay. Vector-borne and Zoonotic diseases. 2007;7(4):617-28.
17. Molaei G, Farajollahi A, Scott JJ, Gaugler R, Andreadis TG. Human bloodfeeding by the recently introduced mosquito, Aedes japonicus japonicus, and public health implications. Journal of the American Mosquito Control Association. 2009;25(2):210-4.
18. Pizarro JC, Lucero D, Stevens L. A method for the identification of guinea pig blood meal in the Chagas disease vector, Triatoma infestans. Kinetoplastid Biology and Disease. 2007;6:1. doi: 1475-9292-6-1 [pii].
19. Pizarro JC, Stevens L. A new method for forensic DNA analysis of the blood meal in Chagas disease vectors demonstrated using Triatoma infestans from Chuquisaca, Bolivia. PLoS One. 2008;3(10):e3585.
20. Stevens L, Monroy MC, Rodas AG, Dorn PL. Hunting, Swimming, and Worshiping: Human Cultural Practices Illuminate the Blood Meal Sources of Cave Dwelling Chagas Vectors (Triatoma dimidiata) in Guatemala and Belize. PLoS Negl Trop Dis. 2014;8(9):e3047.
21. Klotz SA, Schmidt JO, Dorn PL, Ivanyi C, Sullivan KR, Stevens L. Free-roaming kissing bugs, vectors of Chagas disease, feed often on humans in the Southwest. The American Journal of Medicine. 2014;127(5):421-6.
22. Stevens L, Dorn PL, Hobson J, de la Rua NM, Lucero DE, Klotz JH, et al. Vector blood meals and Chagas disease transmission potential, United States. Emerging Infectious Diseases. 2012;18(4):646-9. doi: 10.3201/eid1804.111396 [doi].
23. Peña VH, Fernández GJ, Gómez-Palacio AM, Mejía-Jaramillo AM, Cantillo O, Triana-Chávez O. High-resolution melting (HRM) of the cytochrome B gene: a powerful approach to identify blood-meal sources in Chagas disease Vectors. PLoS Negl Trop Dis. 2012;6(2):e1530.
24. Buitrago NLR, Bosseno MF, Waleckx E, Brémond P, Vidaurre P, Zoveda F, et al. Risk of transmission of Trypanosoma cruzi by wild Triatoma infestans (Hemiptera: Reduviidae) in Bolivia supported by the detection of human blood meals. Infection, Genetics and Evolution. 2013;19:141-4.
25. Buitrago R, Bosseno M-F, Depickère S, Waleckx E, Salas R, Aliaga C, et al. Blood meal sources of wild and domestic Triatoma infestans (Hemiptera:

Reduviidae) in Bolivia: connectivity between cycles of transmission of Trypanosoma cruzi. Parasites & Vectors. 2016;9(1):1.
26. Bosseno MF, Garcia LS, Baunaure F, Gastelum EM, Gutierrez MS, Kasten FL, et al. Identification in triatomine vectors of feeding sources and Trypanosoma cruzi variants by heteroduplex assay and a multiplex miniexon polymerase chain reaction. The American Journal of Tropical Medicine and Hygiene. 2006;74(2):303-5. doi: 74/2/303 [pii].
27. Humair PF, Douet V, Moran Cadenas F, Schouls LM, Van De Pol I, Gern L. Molecular identification of bloodmeal source in Ixodes ricinus ticks using 12S rDNA as a genetic marker. Journal of Medical Entomology. 2007;44(5):869-80.
28. Kent RJ. Molecular methods for arthropod bloodmeal identification and applications to ecological and vector- borne disease studies. Molecular Ecology Resources. 2009;9(1):4-18.
29. Niare S, Berenger J-M, Dieme C, Doumbo O, Raoult D, Parola P, et al. Identification of blood meal sources in the main African malaria mosquito vector by MALDI-TOF MS. Malaria Journal. 2016;15(1):1.
30. Önder Ö, Shao W, Kemps BD, Lam H, Brisson D. Identifying sources of tick blood meals using unidentified tandem mass spectral libraries. Nature Communications. 2013;4:1746.
31. Ngo KA, Kramer LD. Identification of mosquito bloodmeals using polymerase chain reaction (PCR) with order-specific primers. Journal of Medical Entomology. 2003;40(2):215-22.
32. Souza RdCMd, Soares AC, Alves CL, Lorosa ES, Pereira MH, Diotaiuti L. Feeding behavior of Triatoma vitticeps (Reduviidae: Triatominae) in the state of Minas Gerais, Brazil. Memórias do Instituto Oswaldo Cruz. 2011;106(1):16-22.
33. Canals M, Cruzat L, Molina MC, Ferreira A, Cattan PE. Blood host sources of Mepraia spinolai (Heteroptera: Reduviidae), wild vector of chagas disease in Chile. Journal of Medical Entomology. 2001;38(2):303-7.
34. Calzada JE, Pineda V, Garisto JD, Samudio F, Santamaria AM, Saldana A. Human trypanosomiasis in the eastern region of the Panama Province: new endemic areas for Chagas disease. The American Journal of Tropical Medicine and Hygiene. 2010;82(4):580-2. doi: 10.4269/ajtmh.2010.09-0397 [doi].
35. Breniere SF, Pietrokovsky S, Gastelum EM, Bosseno MF, Soto MM, Ouaissi A, et al. Feeding patterns of Triatoma longipennis Usinger (Hemiptera, Reduviidae) in peridomestic habitats of a rural community in Jalisco State, Mexico. Journal of Medical Entomology. 2004;41(6):1015-20.
36. Freitas SPC, Lorosa ES, Rodrigues DCS, Freitas ALC, Gonçalves TCM. Feeding patterns of Triatoma pseudomaculata in the state of Ceara, Brazil. Revista de Saúde Pública. 2005;39(1):27-32.
37. Gonçalves TCM, Rocha DS, Cunha RA. Feeding patterns of Triatoma vitticeps in the State of Rio de Janeiro, Brazil. Revista de Saúde Pública. 2000;34(4):348-52.
38. Vasquez AMd, Samudio FE, Saldaña A, Paz HM, Calzada JE. Eco-epidemiological aspects of Trypanosoma cruzi, Trypanosoma rangeli and their

vector (Rhodnius pallescens) in Panama. Revista do Instituto de Medicina Tropical de São Paulo. 2004;46(4):217-22.
39. Greenwalt DE, Goreva YS, Siljestrom SM, Rose T, Harbach RE. Hemoglobin-derived porphyrins preserved in a Middle Eocene blood-engorged mosquito. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(46):18496-500. doi: 10.1073/pnas.1310885110 [doi].
40. Laskay ÜA, Burg J, Kaleta EJ, Vilcins I-ME, Telford Iii SR, Barbour AG, et al. Development of a host blood meal database: de novo sequencing of hemoglobin from nine small mammals using mass spectrometry. 2012.
41. Laskay ÜA, Breci L, Vilcins I-ME, Dietrich G, Barbour AG, Piesman J, et al. Survival of host blood proteins in Ixodes scapularis (Acari: Ixodidae) ticks: a time course study. Journal of Medical Entomology. 2013;50(6):1282-90.
42. Song Y, Laskay ÜA, Vilcins I-ME, Barbour AG, Wysocki VH. Top-Down-Assisted Bottom-Up Method for Homologous Protein Sequencing: Hemoglobin from 33 Bird Species. Journal of the American Society for Mass Spectrometry. 2015;26(11):1875-84.
43. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2009;37(Database issue):D26-31. Epub 2008/10/23. doi: 10.1093/nar/gkn723. PubMed PMID: 18940867; PubMed Central PMCID: PMCPMC2686462.
44. Mudhasani R, Zhu Z, Hutvagner G, Eischen CM, Lyle S, Hall LL, et al. Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells. The Journal of Cell Biology. 2008;181(7):1055-63. doi: 10.1083/jcb.200802105 [doi].
45. Lian JB, Stein GS, Van Wijnen AJ, Stein JL, Hassan MQ, Gaur T, et al. MicroRNA control of bone formation and homeostasis. Nature Reviews Endocrinology. 2012;8(4):212-27.
46. Ballif BA, Carey GR, Sunyaev SR, Gygi SP. Large-Scale Identification and Evolution Indexing of Tyrosine Phosphorylation Sites from Murine Brain. Journal of Proteome Research. 2008;7(1):311-8. doi: 10.1021/pr0701254.
47. Natarajan C, Hoffmann FG, Lanier HC, Wolf CJ, Cheviron ZA, Spangler ML, et al. Intraspecific polymorphism, interspecific divergence, and the origins of function-altering mutations in deer mouse hemoglobin. Molecular Biology and Evolution. 2015;32(4):978-97. doi: 10.1093/molbev/msu403 [doi].
48. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(12):6940-5. doi: 10.1073/pnas.0832254100 [doi].
49. Goswami T, Li X, Smith AM, Luderowski EM, Vincent JJ, Rush J, et al. Comparative phosphoproteomic analysis of neonatal and adult murine brain. Proteomics 2012;12(13):2185-9. doi: 10.1002/pmic.201200003.
50. Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. Nature Reviews Molecular Cell Biology. 2004;5(9):699-711.

51. Elias JE, Gygi SP. Target-decoy search strategy for mass spectrometry-based proteomics. Proteome Bioinformatics. 2010:55-71.
52. Goloborodko AA, Levitsky LI, Ivanov MV, Gorshkov MV. Pyteomics—a Python framework for exploratory data analysis and rapid software prototyping in proteomics. Journal of The American Society for Mass Spectrometry. 2013;24(2):301-4.
53. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2009;37(Database issue):D5-15. Epub 2008/10/23. doi: 10.1093/nar/gkn741. PubMed PMID: 18940862; PubMed Central PMCID: PMCPMC2686545.
54. Lucero DE, Ribera W, Pizarro JC, Plaza C, Gordon LW, Peña Jr R, et al. Sources of Blood Meals of Sylvatic Triatoma guasayana near Zurima, Bolivia, Assayed with qPCR and 12S Cloning. PLOS Negl Trop Dis. 2014;8(12):e3365.
55. Karlsson AO, Holmlund G. Identification of mammal species using species-specific DNA pyrosequencing. Forensic Science International. 2007;173(1):16-20.
56. Kitano T, Umetsu K, Tian W, Osawa M. Two universal primer sets for species identification among vertebrates. International Journal of Legal Medicine. 2007;121(5):423-7.
57. Melton T, Holland C. Routine forensic use of the mitochondrial 12S ribosomal RNA gene for species identification. Journal of Forensic Sciences. 2007;52(6):1305-7.
58. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403-10.
59. Pinto J, Roellig DM, Gilman RH, Calderón M, Bartra C, Salazar R, et al. Temporal differences in blood meal detection from the midguts of Triatoma infestans. Revista do Instituto de Medicina Tropical de São Paulo. 2012;54:83-8.
60. Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. Electrophoresis 2010;31(11):1764-72. doi: 10.1002/elps.201000027.
61. Johnson RS, Martin SA, Biemann K, Stults JT, Watson JT. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Analytical Chemistry. 1987;59(21):2621-5.
62. Storz JF, Runck AM, Sabatino SJ, Kelly JK, Ferrand N, Moriyama H, et al. Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. Proceedings of the National Academy of Sciences. 2009;106(34):14450-5.
63. Buitrago R, Depickère S, Bosseno M-F, Patzi ES, Waleckx E, Salas R, et al. Combination of cytochrome b heteroduplex-assay and sequencing for identification of triatomine blood meals. Infection, Genetics and Evolution. 2012;12(1):21-7.

64. Gottdenker NL, Chaves LF, Calzada JE, Saldaña A, Carroll CR. Host Life History Strategy, Species Diversity, and Habitat Influence Trypanosoma cruzi Vector Infection in Changing Landscapes. PLOS Neglected Tropical Diseases. 2012;6(11):e1884. doi: 10.1371/journal.pntd.0001884.
65. Ramírez JD, Turriago B, Tapia-Calle G, Guhl F. Understanding the role of dogs (Canis lupus familiaris) in the transmission dynamics of Trypanosoma cruzi genotypes in Colombia. Veterinary Parasitology. 2013;196(1):216-9.
66. Gürtler RE, Cardinal M. Reservoir host competence and the role of domestic and commensal hosts in the transmission of Trypanosoma cruzi. Acta tropica. 2015;151:32-50.
67. Pineda V, Saldaña A, Monfante I, Santamaría A, Gottdenker N, Yabsley M, et al. Prevalence of trypanosome infections in dogs from Chagas disease endemic regions in Panama, Central America. Veterinary Parasitology. 2011;178(3):360-3.
68. Gürtler R, Cecere M, Lauricella M, Cardinal M, Kitron U, Cohen J. Domestic dogs and cats as sources of Trypanosoma cruzi infection in rural northwestern Argentina. Parasitology. 2007;134(01):69-82.
69. Gürtler RE, Cohen JE, Cecere MC, Lauricella MA, Chuit R, Segura EL. Influence of humans and domestic animals on the household prevalence of Trypanosoma cruzi in Triatoma infestans populations in northwest Argentina. The American Journal of Tropical Medicine and Hygiene. 1998;58(6):748-58.
70. Gürtler R, Cécere M, Rubel D, Petersen R, Schweigmann N, Lauricella M, et al. Chagas disease in north-west Argentina: infected dogs as a risk factor for the domestic transmission of Trypanosoma cruzi. Transactions of the Royal Society of Tropical Medicine and Hygiene. 1991;85(6):741-5.
71. Gürtler RE, Cecere MC, Castanera MB, Canale D, Lauricella MA, Chuit R, et al. Probability of infection with Trypanosoma cruzi of the vector Triatoma infestans fed on infected humans and dogs in northwest Argentina. American Journal of Tropical Medicine and Hygiene. 1996;55(1):24-31.
72. Kjos SA, Marcet PL, Yabsley MJ, Kitron U, Snowden KF, Logan KS, et al. Identification of bloodmeal sources and Trypanosoma cruzi infection in triatomine bugs (Hemiptera: Reduviidae) from residential settings in Texas, the United States. Journal of Medical Entomology. 2013;50(5):1126-39.
73. Sasaki H, Rosales R, Tabaru Y. Host feeding profiles of Rhodnius prolixus and Triatoma dimidiata in Guatemala (Hemiptera: Reduviidae: Triatominae). Medical Entomology and Zoology. 2003;54(3):283-9.
74. Wickramasekara S, Bunikis J, Wysocki V, Barbour AG. Identification of residual blood proteins in ticks by mass spectrometry proteomics. Emerging Infectious Diseases. 2008;14(8).

CHAPTER 3: PROTEIN MASS SPECTROMETRY EXTENDS TEMPORAL BLOOD MEAL DETECTION OVER POLYMERASE CHAIN REACTION IN MOUSE-FED CHAGAS DISEASE VECTORS

Judith I. Keller[1], Bryan A. Ballif [1,*], Riley M. St. Clair[1], James J. Vincent[1], M. Carlota Monroy [2,1], and Lori Stevens [1,*]

[1] Department of Biology, University of Vermont, Burlington, Vermont, United States of America

[2] Southwestern Biological Institute, Tucson, Arizona, United States of America

*Co-corresponding authors:

Department of Biology, University of Vermont, Burlington, Vermont, United States of America

Email: lori.stevens@uvm.edu (LS), bballif@uvm.edu (BAB)

3.1 Abstract

Background: Chagas disease is highly prevalent in Latin America, and vector control is the most effective control strategy to date. We have previously shown that liquid chromatography tandem mass spectrometry (LC-MS/MS) is a valuable tool for identifying triatomine vector blood meals.

Objectives: The purpose of this study was to determine blood meal detection ability as a function of method (PCR vs. LC-MS/MS), time since feeding, and the effect of molting in mouse-fed triatomine insect vectors targeting hemoglobin and albumin proteins with LC-MS/MS and SINE-based PCR.

Methods: We experimentally fed *Triatoma protracta* on mice and used LC-MS/MS to detect hemoglobin and albumin peptides over time post-feeding and post-molting (≤12 weeks). We compared LC-MS/MS results with those of a standard PCR method based on Short-Interspersed-Nuclear-Elements.

Findings: Hemoglobin-based LC-MS/MS detected blood meals most robustly at all time points post-feeding. Post-molting, no blood meals were detected with PCR, whereas LC-MS/MS detected mouse hemoglobin and albumin up to 12 weeks.

Main Conclusions: In our study, the hemoglobin signature in the insect abdomen lasted longer than that of albumin and DNA. LC-MS/MS using hemoglobin shows promise for identifying triatomine blood meals over long temporal scales and even post-molting. Clarifying the frequency of blood-feeding on different hosts can foster our understanding of vector behavior and may help devise sounder disease-

control strategies, including Ecohealth (community based ecosystem management) approaches.

3.2 Introduction

Chagas disease is a major neglected tropical disease, with high endemic prevalence in Latin America where most transmission is by insect vectors in the subfamily Triatominae, also known as kissing bugs. Although autochthonous cases of Chagas disease in the United States are thought to be rare, Chagas incidence is likely, for a variety of reasons, to be underreported [1-3]. Chagas disease is caused by the protozoan parasite *Trypanosoma cruzi*, with the most common mode of transmission via a triatomine insect defecating while taking a blood meal or through the oral transmission route. The parasite is subsequently introduced into the new host's blood stream through a break in the skin or mucous membrane.

Identifying the blood meal sources and feeding patterns of native insect vectors provides valuable data for understanding the ecology and behavior of the vector, presenting the need for research aimed at developing better blood meal identification methods. In addition, blood meal source prevalence can help elucidate local transmission cycles and can provide data for evidence-based vector control strategies [4,5]. Insecticide spraying is often effective in the short term, but reinfestation of insect vectors and the financial burden of frequent large-scale spraying, as well as pyrethroid resistance make this not the most strategic option for battling Chagas disease [6-8]. While massive campaigns of indoor residual insecticide spraying have effectively

reduced introduced vectors in many regions, such as *Rhodnius prolixus* in Central America and *Triatoma infestans* in parts of South America [7,8], for native vectors such as *Triatoma dimidiata* in Central America, a recent study suggests a control strategy for Chagas disease includes a holistic Ecohealth approach that focuses on community participation, education, and vector control [5,9-12]. As such, within the Ecohealth framework, vector control policies need to consider domestic vectors and perhaps underestimate sylvatic vector influence [13]. To develop effective native vector control strategies, a knowledge of the vector ecology, including blood meal sources of vectors collected in various ecotopes, such as sylvatic, peridomestic, or domestic environments is important [5].

However, blood meal source detection can be challenging. Triatomine insects are hematophagous and their digestive systems evolved to digest blood [14]. Therefore, blood meal detection ability rapidly decays as the blood is digested. Few studies examined detection ability over time and have shown all detection methods work best with high quality material from recently-fed vectors [15-19]. Often fieldwork conditions are not ideal for the storage and transportation of insect samples to maintain high quality of DNA or antigens commonly used for blood meal identification. Thus, time since feeding and storage of material could be two of the major reasons why in many studies upwards of 50% of samples do not have a blood meal detected [19-23].

Blood meal sources have traditionally been detected by a number of methods including immunological approaches such as precipitin tests and ELISA [24] which are dependent on protein antibodies of blood meal sources present in an area, and DNA-

based methods often focusing on PCR amplification of mitochondrial or nuclear DNA. In addition, PCR amplification of a species-specific repetitive sequences of nuclear DNA, such as short interspersed nuclear elements (SINEs) using standard PCR [17,18], PCR amplification followed by sequencing of vertebrate mitochondrial DNA [25] or ribosomal subunits such as 12 S [26], genomics [27], and other next-generation sequencing tools [28] are emerging. Recently, we have shown the usefulness of a protein-based method based on liquid chromatography tandem mass spectrometry (LC-MS/MS) for blood meal source identification [29]. A few studies have compared different methodologies directly. For example, Lucero et al. compared 12 S sequencing and qPCR [30], while Stevens et al. compared 12 S sequencing and cytochrome b [25]. We compared LC-MS/MS with 12 S-based DNA sequencing and found LC-MS/MS identified blood meals from insects collected alive and dead; however, blood meals were not detected with DNA sequencing methods for three of the four samples [29]. Blood meal detection ability can also vary depending on the species of host blood. A study examining the effect of both blood meal host species and time elapsed since a recent feeding, reported detection ability can drop off as early as 1-2 weeks for some species [18].

Blood is a complex fluid with components that decay at different rates. Because iron stabilizes molecules, blood proteins such as hemoglobin can be remarkably stable and have been detected in a 46-million-year-old fossilized mosquito [16,31,32]. Peptides from hemoglobin, the most abundant protein in red blood cells, have been detected 309 days post-molt in ticks; and peptides from albumin, the most abundant blood serum

protein, have been detected 85 days post-molt in ticks [16]. Other components such as transferrin and immunoglobulins are known to degrade rapidly, while keratin, actin, histones, and tubulins were too conserved across the animal kingdom to be informative when evaluated for mass-spectrometry-based detection in ticks [16]. Abundance, stability, and species-specific sequence variation make hemoglobin and albumin molecules great targets for LC-MS/MS-based techniques [16,29,32]. LC-MS/MS has led to promising results when compared with DNA-based methods [29]. A major benefit of using protein-based techniques is the quantity and quality of data gained from a single LC-MS/MS run [16].

Information about blood meal perseverance in the insect gut is limited. Nevertheless, the longevity of blood meal detection ability in Chagas vectors is critical information to aid in the knowledge of the parasite transmission and overall feeding habits of the insect. For example, the average time for a blood meal digestion has been estimated to be approximately 14 days in adult female *Triatoma infestans* [33], but different components of blood may vary and blood meals have been detected up to 10 weeks post-feeding in adult male and female *T. pallidipennis*, *T. barberi*, *T. dimidiata, T. phyllosoma,* and *T. longipennis* [15]. Molting behavior of insects can also affect blood meal detection. Kissing bugs are hemimetabolous, emerging wingless from an egg, and successively molt through five nymphal instars into winged adults [34]. Experimentally evaluating blood meal detection post-molt would indicate whether or not we can detect feeding across molts and if the decrease in albumin and hemoglobin peptides over time can provide temporal information about the last blood meal.

In this study we determine how the ability to detect and identify a blood meal declines over time and is affected by molting. We assess Triatominae insect vector blood meals for the first time comparing two identification methods, protein LC-MS/MS and PCR of SINE-DNA in two experiments: (1) recently molted adult *Triatoma protracta* (Hemiptera: Reduviidae) fed once on mouse and assayed 0-4 weeks post-feeding; and (2) recently molted (fed approximately 1 week prior) *T. protracta* not fed after eclosion and assayed 0-12 weeks after molting. In addition to comparing LC-MS/MS with PCR, within LC-MS/MS we compare the ability of hemoglobin and albumin peptides to identify the source of a blood meal.

3.3 Methods

3.3.1 Ethics

White inbred ICR (CD-1) *Mus musculus* (house mouse) (Harlan Laboratories, Madison, WI) were used for feeding experiments. All procedures using mice were first approved by the Southwestern Biological Institute, Tucson, Arizona, USA Animal Care and Use Committee and follow international standards [(35)]. Mice were immobilized in small mesh cages and placed in the enclosure containing the insect vectors. Insects were allowed to feed until satiated, or approximately for 30-60 minutes, at which point mice were removed.

3.3.2 Experimental Feeding and Insect Collection

The parent colony of *Triatoma protracta* vectors used in this study was established in 2009 from wild caught vectors collected in the Tucson, Arizona, USA

basin with approximately one dozen individuals, and no additions have been made since then. As a result, we consider this population to be relatively inbred and individuals have been in colony for 7 generations. The insects were housed in 30 x 25 x 28 cm polypropylene containers containing a paper lining at the base and egg cartons (composed of wood pulp fiber) for refugia. They were provided a continuous source of water, which consisted of 1.5 mL microtubes filled with distilled water and plugged with absorbent cotton. The typical temperature in the rearing facility ranged from approximately 10-20º C in the winter and 20-32º C in the summer, with a relative humidity of approximately 30 % in the winter, and 10-40 % in the summer. Spring and fall temperatures and humidity were intermediate between those of winter and summer. In March-June 2016, the lab-reared *T. protracta* were established into two groups representing the two experiments: (1) post-feeding or (2) post-molting. After the experimental mouse feeding, the two experimental groups were kept separately in equivalent aforementioned containers. No non-experimental insects were housed in the experimental groups.

The post-feeding (F) insects were collected after eclosing as adults and allowed a single blood meal on *M. musculus* within 19-37 days. Within an hour of feeding, the adult individuals (hereafter referred to as F0wk) were preserved in 95 % ethanol and 5 % glycerol and stored at 4 ºC. Additional individuals were collected at each of the following time points: 1 week (F1wk), 2 weeks (F2wk), and 4 weeks (F4wk) post-feeding without access to an additional blood meal. Insects that died between sampling times were not analyzed (see Supplementary Table III for details of sample sizes and

longevity). None of the insects in the F group survived past 4 weeks, except one specimen that was not analyzed due to a mishap in the preparation. We analyzed four insects from each time period, except F0wk where we analyzed three.

Adult post-molting (M) insects were collected from the colony after feeding on *M. musculus* as 5th instar nymphs, and most molted to adults within 1 week of feeding. After molting, insects in this group were not fed, but were collected and preserved at the same time intervals (M0wk, etc) as the fed insects. These post-molt insects survived longer than the fed insects and sampling was extended to 8 and 12 weeks (Table 1). The longevity of post-molt bugs was a little longer, of the 9 remaining at 8 weeks, 8 died before week 12 and we were able to analyze the one specimen alive at 12 weeks (Supplementary Table III). Although most adult *Triatoma* species live much longer than 4-8 weeks, starvation has been shown to reduce adult longevity in *Rhodnius prolixus* [(36)], and the difference between the post-molt and post-fed specimen may be an artifact of small sample size. We sampled two insects from each time period, 0, 1, 2, 4, 8 and one surviving individual at 12 weeks. Samples were stored at 4º C and within 1-4 weeks of collection, were shipped to the University of Vermont by priority mail in insulated containers, where they were stored at -20 ºC until dissection in August 2016 and May 2017.

The 15 post-feeding and 11 post-molting *T. protracta* were evaluated using methods similar to Keller et al. [(29)]. Results were visualized with graphs made using JMP®, Version 13 (SAS Institute Inc., Cary, NC, 1989-2016). Below we highlight important aspects of the methods and indicate changes from our previous study [(29)].

3.3.3 Dissection of Insect Vectors

For each insect, the abdomen was cut into left and right halves. Abdomen halves were randomly assigned to LC-MS/MS protein analysis or mouse-specific SINE-DNA PCR.

3.3.4 Hemoglobin and albumin protein extraction, SDS-PAGE, and mass spectrometry

We extracted protein from *T. protracta* insect abdomen halves as previously described [29] except that 200 µl of denaturing sampling buffer was added per 0.1 g of insect tissue. For samples weighing 0.05 g and below, 100 µl of 95 ˚C denaturing sampling buffer was added. Denaturing SDS-PAGE using gel regions surrounding the molecular weight of hemoglobin (~16 kDa) and albumin (~65 kDa) were excised and prepared for mass spectrometry analysis as previously described [29]. In brief, because LC-MS/MS works on peptides smaller than the hemoglobin and albumin proteins, following in-gel digestion with trypsin and peptide extraction, LC-MS/MS was performed using a linear ion trap-orbitrap (LTQ-Orbitrap; Thermo Electron, Waltham, Massachusetts, USA) where spectra all were collected in the orbitrap [29,37]. Samples were subjected to 15 min of isocratic loading in 2.5 % MeCN, 0.15 % FA (Solvent A), and peptides were subsequently eluded with a 0-50 % gradient of 99 % MeCN, 0.15 % FA (Solvent B) over 45 min (400 nl/min flow rate average across a flow splitter), followed by 10 min 100 % Solvent B, and a 15 min equilibration with Solvent A.

LC-MS/MS does not directly sequence peptides, but rather infers amino acid sequences of short peptides based on theoretical peptide masses present in an

underlying database, in our case, GenBank [38] hemoglobin and albumin entries. We searched these mass spectra using the SEQUEST algorithm (Thermo Electron V26.12) against a custom forward and reverse concatenated database containing vertebrate hemoglobin sequences (20 January 2016, 17,000+ entries) extracted from GenBank as previously described [29] and "serum albumin" (26 October 2016, 1600+ entries) in any curated field. Peptide identification and stringent filtering of peptides was as described in Keller et al. [29], where no reverse database matches resulted and false discovery of peptides was below 0.01 %.

Blood meal sources were identified as previously described with a pipeline to infer the most likely blood source [29] (Figure 1). The pipeline considered the potential taxa represented by the peptide inferred from the mass spectra and cross-referenced the most likely blood meal source based on individual peptides in a sample. We subsequently quantified the identified protein coverage at the peptide and amino acid level. This provided the percentage support for a particular blood source (see Keller et al. [29] for details).

3.3.5 DNA extraction and SINE-based PCR

DNA extraction used the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) as previously described [17,30]. Briefly, the manufacturer's instructions for extracting tissue were followed using insect abdomens chopped finely with scissors. DNA was eluted with 2 separate sequential elutions of 100 µL each. DNA concentration was measured using a Nanodrop ND-1000 instrument (Thermo Scientific, Waltham, MA,

USA), and the instrument was calibrated using the elution buffer (Qiagen, Valencia, CA) DNA was stored in.

DNA extracts were subjected to PCR amplification after optimizing previously published methods for our equipment and reagents [18,39,40]. The 12 µL PCR reaction contained 1 µL of DNA template, 0.2 µM each of forward (5'AGATGGCTCAGTGGGTAAAGG3') and reverse (5'GTGGAGGTCAGAGGACAAACTT3') primers, and 6 µl 1X EconoTaq PLUS GREEN (Lucigen, Middleton, WI, USA). PCR conditions were as follows: initial denaturization for 5 min at 95 ºC, 30 cycles of 95 ºC for 30 sec, annealing for 30 sec at 55 ºC, and extension for 30 sec at 72 ºC, followed by a final extension of 5 min at 72 ºC. Nancy-520 stained 1.5 % agarose gels (Sigma-Aldrich, Milwaukee, WI, USA) were used to verify the 118 bp PCR fragments. Positive (DNA extracted from mouse tissue) and negative (PCR-grade water) controls were included in each set of PCR amplifications.

3.4 Results

The purpose of this study was to determine blood meal detection ability as a function of method (PCR vs. LC-MS/MS), time since feeding, and the effect of molting in Triatominae insect vectors targeting hemoglobin and albumin proteins with LC-MS/MS and SINE-based PCR. This is the first study to compare protein-based and DNA-based detection methods of experimentally-fed arthropod vectors. Briefly, hemoglobin-based LC-MS/MS yielded the most robust mouse blood meal detection and

identification over time. In our experiments none of the post-feeding specimens lived longer than 4 weeks; however, we were able to identify the *M. musculus* blood meals based on hemoglobin and albumin peptides 4 and 2 weeks post-feeding, respectively (Figure 2 and Figure 3, Supplementary Figure 2). With SINE-based PCR we were able to detect *M. musculus* blood meals up to 1 week post-feeding. The post-molting triatomines lived longer, and we were able to identify *M. musculus* hemoglobin and albumin peptides both up to 12 weeks post-molting, while SINE-based PCR detected no *M. musculus* blood meals at any time post-molting (Figure 3, Supplementary Figure 2).

3.4.1 Post-feeding

LC-MS/MS- hemoglobin: Using LC-MS/MS we were able to identify hemoglobin peptides throughout the entire sampling time of 4 weeks post-feeding, except for one F4wk sample; however the number of hemoglobin peptides decreased over time from an average of over 300 at F0wk to less than 10 at F4wk (Figure 2, Table I). For all four of the biological replicates at the 0, 1, and 2 wk time points and one of the F4wk samples the combination of peptides unambiguously identified the blood meal to the species level, *M. musculus*. With the other three F4wk samples, for two we were able to identify 2-6 species as the most probable blood source (which in both cases included the correct *M. musculus* blood meal - Sample ID 49 equally supported 6 species: *M. musculus*, *Otospermophilus beecheyi*, *Mus spretus*, *Mus minutoides*, *Jaculus jaculus*, *Callospermophilus lateralis*; Sample ID 51 equally supported *M.*

musculus and *M. spretus*), while as stated above, one replicate did not contain hemoglobin peptides (Table I).

LC-MS/MS- albumin: We were less successful detecting albumin peptides over time. We were only able to identify albumin peptides up to 2 weeks post-feeding and like hemoglobin, the number decreased over time from an average of over 100 at F0wk to 0 at F4wk (Figure 2, Table II). For all the replicates at F0wk and F1wk, and three of the four replicates at F2wk, we identified the blood meal to the species level (Figure 3). For the other F2wk sample, there was equally strong support for *Rattus norvegicus* and *M. musculus* as the blood meal source (Table II). Overall, hemoglobin and albumin peptide abundance varied significantly between each time point post-feeding (Least Squares Regression, $p<0.001$), and albumin was significantly lower in abundance than hemoglobin (Least Squares Regression, $p<0.001$) (Table III).

SINE-DNA: The *M. musculus*-specific-SINE based PCR was the least successful in detecting the *M. musculus* blood meal, we were only able to detect the blood meal in the F0wk and F1wk samples (Figure 3).

3.4.2 Post-molting

LC-MS/MS- hemoglobin: Using LC-MS/MS, we were able to identify hemoglobin peptides throughout the entire sampling time of 12 weeks post-molting, except for one M8wk samples. As expected, for the M0, 1, 2wk specimens, the number of hemoglobin peptides was lower than for the post-feeding experiment and generally decreased over time from an average of over 100 at M0wk to less than 10 at M8wk (Figure 2, Table I). Spectral counts were more variable in post-molting individuals. For

example, M8wk averaged around 8 hemoglobin peptides while the single replicate M12wk contained 95 hemoglobin peptides. For all replicates except one M8wk sample we identified the blood meal to the species level. One M8wk sample contained a single hemoglobin peptide matching 261 species (Table I).

LC-MS/MS- albumin: As with the post-feeding specimens, we were less successful detecting albumin peptides post-molting. We were able to detect albumin peptides over all time points, but the number decreased from a range of 18-49 at the early times to 0-3 at the later (Figure 2, Table II). For all the replicates at 0, 1, and 2 wk and one of the two M4wk, we identified the blood meal to the species level. For the other samples, we narrowed the likely blood meal sources to 2-7 species, again, all including known blood meal source *M. musculus* (Sample ID 54 equally supported *M. musculus* and *R. norvegicus*; Sample ID 23 and 30 equally supported *M. musculus, R. norvegicus, Sorex araneus, Octodon degus, Ochotona princeps, Dipodomys ordii,* and *Cricetulus griseus*.). Albumin peptides appeared to decrease significantly after the 2 week timepoint (Table II). Overall, hemoglobin and albumin peptide abundance did not vary significantly post-molting (Least Squares Regression, $p>0.05$) (Table III).

SINE-DNA: Our *M. musculus*-specific SINE-based PCR did not detect *M. musculus* blood meals at any time points post-molting (Figure 3).

3.5 Discussion

Ideally, a combination of assays gives researchers a diverse toolbox for identification of blood meal sources. For blood-feeding arthropod disease vectors,

including the triatomine vectors of Chagas disease, increasing our understanding of vector blood meal sources facilitates the design of evidence-based control strategies. Few papers have compared different blood meal detection techniques directly, e.g. [30,31], and this is the first paper to compare a LC-MS/MS protein-based approach to a well-established DNA-based approach. The assay based on protein identified the blood meal source for a longer time post-feeding as well as post-molting.

This study shows variation in detection of blood meals by LC-MS/MS and mouse-specific SINE-based PCR as a function of time since last feeding and time since molting. Overall, LC-MS/MS based on hemoglobin gave precise blood meal identification for the longest amount of time post-feeding and post-molting. Albumin peptides were also present for the majority of time points in both experiments, however, detection of albumin peptides dropped off sharply 2 weeks post-feeding. DNA-based detection using mouse-specific SINE-DNA was only successful in detecting fresh blood meals up to 1 week post-feeding and did not detect a blood meal at any time points for the post-molt experiment. To our knowledge, this study is the first to explore post-molt blood meal detection in triatomines, as only incidental findings have been reported previously [19].

Using hemoglobin peptides opens the door to exploring blood sources of early stage nymphs that have molted into later stages and possibly adults. Many studies of blood meal sources, e.g. [41], focus on later stage triatomines, which are more likely to be encountered and are more mobile. Because we were able to detect peptides from a previous life stage up to 12 weeks post-molting with LC-MS/MS but not DNA, we

suggest further studies that explore blood sources of early stage nymphs. Such future studies could include examining the nymphs themselves, or through experimental feeding to determine if blood sources of early stage nymphs could be detected in later nymphs or even adults. The SINE-based PCR did not detect any blood meals post-molt. Albumin did have a lingering signature post-molt, but recovered peptides were significantly lower in number. While protein analyses do not have the luxury of amplification technologies that exist for DNA, LC-MS/MS is becoming more sensitive. Because of the low fmol (femtomole = 10^{-15}) amount of material needed for LC-MS/MS, future studies should examine the ability to detect blood proteins over the life cycle of the insect. The development and standardization of methods for LC-MS/MS detection will increase our ability to detect blood sources in field-collected samples, and can boost the design of control strategies for the distinct transmission cycles maintained by different vector species in different localities.

An increasing number of studies are trying mass spectrometry-based techniques for blood meal identification in insect vectors [16,42,43]. Our mass spectrometry-based technique differs from DNA sequencing in that the molecule is not sequenced per se, but rather highly accurate mass measurements are used to match observed mass spectra with theoretical ones based on protein sequences in GenBank [29,32]. An alternative approach, matching spectral libraries using measurements made by Matrix Assisted Laser Desorption/Ionization time of flight mass spectrometry (MALDI-TOF) has also been used [42,43] but requires blood controls from animal species likely encountered in the field to make the matching spectral libraries. However, obtaining such material can

be problematic and precludes identifying unanticipated taxa. LC-MS/MS matching theoretical to observed spectra provides a practical approach using sequence searching with data readily available in GenBank [29], which we only expect to grow in sequence information over time. In addition, as management decisions are often based at higher taxonomic levels (e.g. rodents in general, birds in general), identifying closely related species through conservation of hemoglobin sequences of related species allows for comprehensive vector management using LC-MS/MS data. In our study, the hemoglobin signature in the blood lasted longer than that of albumin, which has also been previously shown [16]. We show that hemoglobin is detected in *T. protracta* even several weeks after the insect had fed, and was detected even longer, notably from fewer peptides, in insects that had molted but not subsequently fed (4 and 12 weeks, respectively).

Given hemoglobin is more abundant in blood and is highly stable, it is not necessarily surprising that it is detectable longer than albumin. However, this might not have been the case, particularly as albumin is also stable and as a larger molecule compared to hemoglobin (~608 vs. ~289 amino acids) offers more total tryptic peptides for identification by LC-MS/MS [38]. Hemoglobin has a second advantage regarding peptide identification in that the size of the underlying database, in this case all hemoglobin entries currently in GenBank, is larger than that for albumin. Albumin has an order of magnitude fewer entries (>17,000 for hemoglobin (20 January 2016) vs. <1,700 for albumin (26 October 2016)), which could be a problem when examining samples from sylvatic vectors collected in regions with little molecular data on

vertebrate biodiversity. Even if less useful for species identification, the presence or absence of albumin peptides could provide an estimate of the time window in which the insect vector fed, although this would require more controlled experiments to develop a range for time window estimates.

Blood meals have only occasionally been detected post-molting in triatomine insect vectors [19], and there is a gap in literature on this subject. Although we used positive controls (DNA extracted from mouse tissue) and checked for PCR inhibition with our samples (internal, same-tube controls), we did not detect a blood meal post-molt. Theoretically PCR can amplify from a single molecule, but it is likely that in the insect vector, DNA from mouse blood was not of sufficient quality or too low in abundance, and PCR requires an intact DNA strand for the sequence between the primers. With DNA we targeted a 118 bp fragment with PCR, however, this SINE transposable element has an estimated 2000 copies in the mouse genome and a detection limit of 0.01 ng 10^{-5} g using qPCR [39]. In contrast, LC-MS/MS easily analyzes small peptides and surveys all peptides extracted from the insect digestive system for a match to hemoglobin peptides reported in GenBank. Experimental feeding studies looking at blood meal detection at times post-feeding by Pinto et al. (2012) had similar results. They were able to detect a mouse blood meal in *Triatoma infestans* 14, but not 21, days after feeding and reported a detection limit of 10 ng with ethidium bromide stained agarose gels, using the same mouse-specific PCR assay used in this study. Hemoglobin-based LC-MS/MS can potentially fill the gap in knowledge of an insect vector's previous blood meal after the insect has molted or not fed for long

periods of time. Triatomine nymphs need at least one blood meal to molt to the subsequent life stage and female vectors generally need a blood meal before egg laying, although autogeny has been recorded in some species of kissing bugs [44].

Triatomine vector life spans vary, but have been recorded to last from several months to over a year [45,46] and feeding patterns can change from various nymphal stages to adult stages, as well as over the life span of an adult [47]. In addition, triatomine species can exhibit opportunistic feeding behaviors related to the relative abundance and proximity of animal blood sources [48]. Therefore, detection of blood meals at various times post-feeding as well as post-molting, and elucidating blood meal sources from previous life stages, is an important aspect of making Ecohealth-based management decisions especially for native vectors.

In this study we show that LC-MS/MS allows correct identification of blood meal sources to the species level but the taxonomic level of resolution decreases with time and as a function of molting and the molecule examined. Although > 95% of amino acids identified were previously reported in GenBank in mouse, not all peptides identified matched the known blood source. Indeed, even with sequencing, a 100% match with DNA is also not always possible because of previously unidentified DNA polymorphisms. We chose to use mouse for our controlled blood meal as mice are easily available and have been used in previous feeding studies [15,18]. However, they also come with challenges such as heterozygosity and various chromosome locations of hemoglobin genes [49].

As we are critically examining the strengths and weaknesses of this LC-MS/MS technique, we spent considerable energy investigating the few instances when the peptide identified was not known to match to the known blood meal source. Two likely explanations for these mis-matches, are: previously unknown polymorphisms, and, misidentification by the SEQUEST program. When our sample pool is corrected for likely cases of misidentification by SEQUEST, our lowest identify of 95.3% amino acids matching mouse increases to 99.7 % (294/295 amino acids). If the same approach is applied to all samples, all blood source identification confidence based on amino acids increase to greater than 99.5 %, and it may be a fair assumption that the 0.5% represent additional unknown polymorphisms. We detail in the supplementary material specific examples of misidentification by SEQUEST (See Supplementary Table I and Supplementary Figure 1). Furthermore, for a field study if several species match as the most likely blood meal source (e.g. Samples 26, 30, Table II), one could easily rule out those whose biogeography does not overlap with the species of Chagas vector examined, e.g., the Eurasian shrew, *Sorex araneus*. For others, e.g., American pika, *Ochotona princeps*, there is equal support from a single specimen for both pika and mouse. One would be able to comment on the likelihood of each based on other specimens examined in the same study. In addition, published host records, e.g. [26], would indicate if the host had been previously reported for Chagas vectors, although novel blood meal sources are regularly detected and need to be considered.

The use of a blood meal detection technique known to accurately detect blood meals across long temporal scales such as LC-MS/MS can lead to a better

understanding of vector biology and for developing informed strategies for vector control. Chagas and other arthropod disease vectors have often not fed recently, but may contain remnants of a blood meal from some time ago. Indeed, in some studies field-collected triatomines were mostly found unfed (*Rhodnius prolixus*) [(50)] and while 5th instar nymphs feed most frequently (*Meccus pallidipennis*) [(51)], nymphs especially are capable of surviving long period of starvation [(50)]. In addition to time since feeding, the quantity of a blood meal is likely to affect detection. Although in our experiment vectors were allowed to feed until satiated, this is not always the case in the wild.

In addition to enhancing our ability to detect blood meals for longer times after feeding, hemoglobin-based LC-MS/MS might also be able to detect multiple blood meals. Wild vectors often have multiple blood meal sources, a topic we have yet to address with our detection technique. As suggested previously [(29)] further studies could examine the possibility of using synthetic peptides such as AQUA (*A*bsolute *QUA*ntification) peptides for quantification of a blood meal. Spiking a synthetic AQUA peptide into a blood meal sample could aid in quantification of a blood meal and allow more detailed detection of multiple blood meals. Our previous study showed the feasibility [(29)] of this approach.

Freshly fed triatomine specimens are ideal for blood meal analysis. Storage conditions may affect detection ability and fieldwork conditions are not perfect for the storage and transportation of insect samples to maintain high quality of DNA or antigens commonly used for blood meal identification [(52)], however, we showed previously (Keller et al. 2017) and in this study that LC-MS/MS seem less sensitive to

the storage condition of the vectors. In addition, detection of non-recent blood meals is important to developing Ecohealth management decisions and developing vector control strategies. Therefore, using a technique such as hemoglobin-based LC-MS/MS has strong advantages in some situations for identifying blood meal sources, such as the ability of a single LC-MS/MS run to identify all blood meals at the same time and is also reasonably priced (see Keller et al. 2017, Supplementary Table 3 and Önder et al. 2013, Supplementary Table S3 for review on cost analysis) and proteomics resources are available in many areas with endemic triatomine populations [53]. LC-MS/MS assays based on hemoglobin and potentially other proteins are a powerful tool for evaluating blood meals in Chagas disease vectors, and could be applied to other vector disease systems. The ability to detect blood proteins over long temporal scales and in molted individuals opens the door to using LC-MS/MS hemoglobin-sequence-based techniques in field-collected specimens and is a valuable part of the diverse toolbox for identification of blood meal sources.

3.6 Acknowledgements

Thank you to Bethany Ahlers for helpful guidance with mass spectrometry analysis and the reviewers for helpful comments.

3.7 Figures and Tables

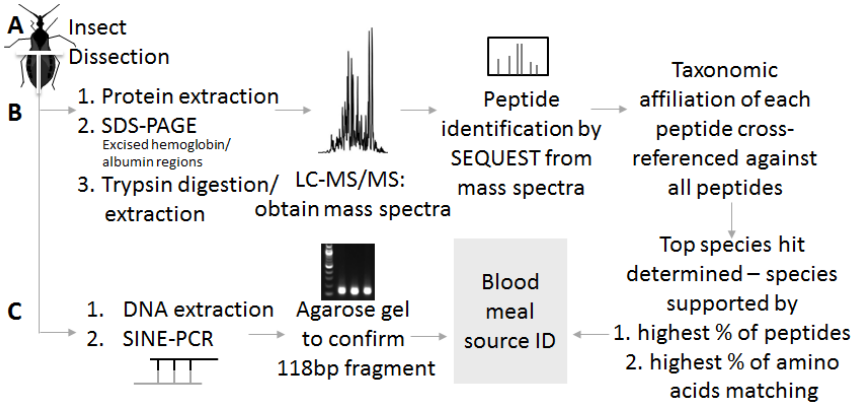

Figure 3.1 Workflow describing LC-MS/MS and DNA-based detection of blood meals. Insect abdomens were dissected into left and right halves (A) and subsequently were processed for blood meal species identification by protein (B) (see Keller et al. 2017 for additional details) and by DNA (C).

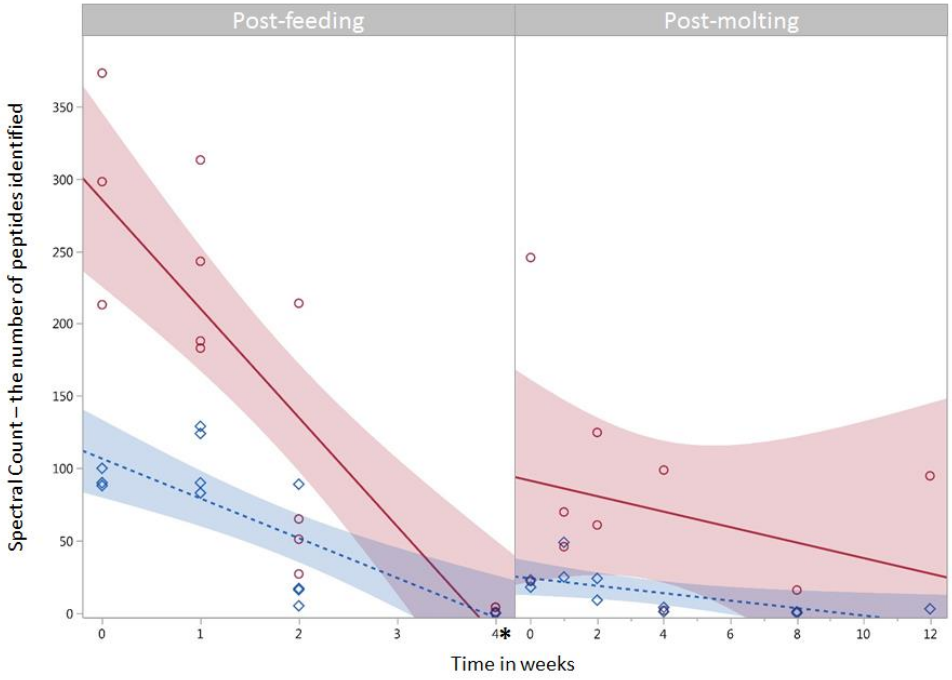

Figure 3.2: Detection of hemoglobin and albumin peptides over time from a known *M. musculus* blood meal source. The number of hemoglobin (red open circle/solid line) and albumin (blue diamond/dotted line) spectral counts decreased over time post-feeding (left) and post-molting (right). There were more peptides at 0 and 2 wk in the post-feeding specimens, but more at 4 wk for the post-molt. Linear regression lines were fit and 95% confidence intervals are shown (shading) with an alpha level of 0.05. Hemoglobin and albumin peptide abundance varied significantly between each time point post-feeding (Least Squares Regression, $p<0.001$), and albumin was significantly lower in abundance than hemoglobin (Least Squares Regression, $p<0.001$). Note the difference in time scales on the x-axis (*), none of the analyzed post-feeding specimen lived longer than 4 weeks.

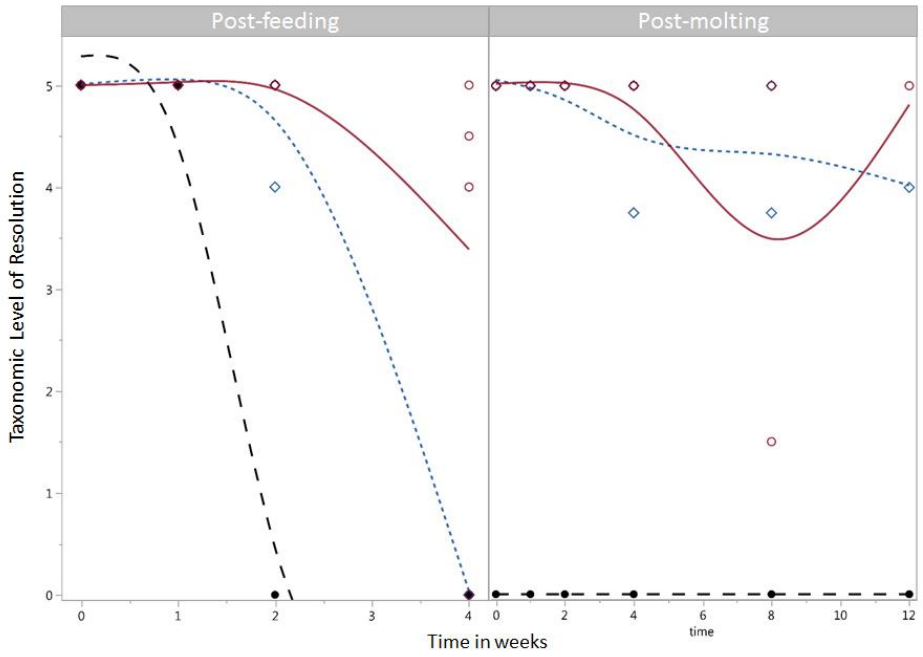

Figure 3.3: Taxonomic level of resolution for *M. musculus* blood meals over time. Taxonomic level of resolution by LC-MS/MS varied between hemoglobin (red open circle/solid line) and albumin (blue diamond/dotted line) proteins and from SINE-based PCR (black closed circle/lined line) post-feeding (left). DNA, albumin, and hemoglobin provided species-specific blood meal identification up to 1, 2, and 4 weeks post-feeding, respectively. Post-molting (right), the taxonomic level of resolution for *M. musculus* blood meals by LC-MS/MS was stronger for hemoglobin and albumin than SINE DNA, which never detected a blood meal at any post-molting time point. A best-fit line (cubic smoothing spline, lambda of 0.0855, standardized X values) and 95% confidence intervals are shown.

Table 3.1: Summary statistics for hemoglobin peptides identified in adult *Triatoma protracta* abdomen halves in post-feeding and post-molting specimens. (AA) amino acid; [a] the number of unique peptides identified in a sample that match the known blood meal source, *M. musculus*, of the approximately 23 detectable peptides from trypsin digestion of 142 aa alpha and 147 aa beta hemoglobin (depending on the amino acid variation; based on GenBank entries NP_032244.2, BAG16710.1); [b] number of amino acids of the unique peptides identified that match mouse; [c] spectral count, or number of hemoglobin peptides identified in each LC-MS/MS run; [d] *M. musculus* was not uniquely identified as the most likely blood meal source. Sample ID 49 equally supported 6 species: *M. musculus*, *Otospermophilus beecheyi*, *Mus spretus*, *Mus minutoides*, *Jaculus jaculus*, *Callospermophilus lateralis*; Sample ID 51 equally supported *M. musculus* and *M. spretus*; The single peptide from Sample ID 27 was a non-specific that had been reported in *M. musculus* and 260 other species.

			Mus musculus match						
	Sample ID		**no. unique peptides match/total unique[a]**	**% peptides matching *M. musculus***	**Total AA identified based on no. of unique peptides[b]**	**no. AA mis-matches**	**% amino acid matching to *M. musculus***	**Spectral Count[c]**	**Average Spectral Count**
1. Post-Feeding	F0wk	33	16/20	80.00	299	8	97.3	298	294.7
	F0wk	34	15/18	83.33	292	9	96.9	373	
	F0wk	35	18/24	75.00	399	16	96.0	213	
	F1wk	38	16/20	85.00	311	12	96.1	188	231.8
	F1wk	40	20/26	76.92	390	12	96.9	183	
	F1wk	41	15/19	78.95	300	8	97.3	243	
	F1wk	42	18/21	85.71	340	9	97.4	313	
	F2wk	43	14/15	93.33	213	1	99.5	65	89.3
	F2wk	44	13/14	92.86	200	1	99.5	51	
	F2wk	45	17/22	77.27	314	9	97.1	214	
	F2wk	46	4/4	100.00	56	0	100.0	27	
	F4wk	49	1/1[d]	100.00	13	0	100.0	1	2.3
	F4wk	50	4/4	100.00	56	0	100.0	4	
	F4wk	51	4/4[d]	100.00	53	0	100.0	4	
	F4wk	52	*no peptides identified*					0	
2. Post-Molting	M0wk	15	23/28	82.14	422	15	96.4	246	134.0
	M0wk	4	10/12	83.33	186	6	96.8	22	
	M1wk	11	19/22	86.36	320	9	97.2	70	58.0
	M1wk	10	19/21	90.48	326	9	97.2	46	
	M2wk	17	20/22	90.91	293	6	98.0	126	93.5
	M2wk	19	10/12	83.33	197	5	97.5	61	
	M4wk	24	17/21	80.95	295	14	95.3	99	50.5
	M4wk	23	2/2	100.00	31	0	100.0	2	
	M8wk	27	1/1[d]	100.00	0	0	100.0	1	8.5
	M8wk	30	8/10	80.00	151	6	96.0	16	
	M12wk	54	18/21	85.71	317	9	97.2	95	95.0

Table 3.2: Summary statistics for albumin peptides identified in adult *Triatoma protracta* abdomen halves in post-feeding and post-molting specimen. (AA) amino acid; [a] Of the approximately 58 detectable peptides from trypsin digestion of the 609 AA albumin protein (depending on AA variation, based on GenBank sequence CAD29888.1), shown are the number of unique peptides identified in a sample that match the known blood meal source, *M. musculus;* [b] number of amino acids of the unique peptides identified that match mouse; [c] spectral count, or number of hemoglobin peptides identified in each LC-MS/MS run; [d] *M. musculus* was not uniquely identified as the most likely blood meal source. Sample ID 46 and 54 equally supported *M. musculus* and *R. norvegicus*; Sample ID 23 and 30 equally supported *M. musculus, R. norvegicus, Sorex araneus, Octodon degus, Ochotona princeps, Dipodomys ordii*, and *Cricetulus griseus*

			Mus musculus match						
	Sample ID		**no. unique peptides match/total unique[a]**	**% peptides matching *M. musculus***	**Total AA identified based on no. of unique peptides[b]**	**no. AA mis-matches**	**% amino acid matching to *M. musculus***	**Spectral Count[c]**	**Average Spectral Count**
1. Post-Feeding	F0wk	33	14/14	100.00	194	0	100.0	90	95
	F0wk	34	14/14	100.00	176	0	100.0	102	
	F0wk	35	14/14	100.00	166	0	100.0	93	
	F1wk	38	8/8	100.00	100	0	100.0	90	108.25
	F1wk	40	5/5	100.00	66	0	100.0	83	
	F1wk	41	15/15	100.00	182	0	100.0	129	
	F1wk	42	15/15	100.00	189	0	100.0	131	
	F2wk	43	9/9	100.00	116	0	100.0	17	30.25
	F2wk	44	9/9	100.00	116	0	100.0	17	
	F2wk	45	12/12	100.00	181	0	100.0	81	
	F2wk	46	3/3[d]	100.00	40	0	100.0	6	
	F4wk	49	*no peptides identified*						0
	F4wk	50	*no peptides identified*						
	F4wk	51	*no peptides identified*						
	F4wk	52	*no peptides identified*						
2. Post-Molting	M0wk	15	14/14	100.00	172	0	100.0	23	20.5
	M0wk	4	11/11	100.00	138	0	100.0	18	
	M1wk	11	19/19	100.00	248	0	100.0	49	37
	M1wk	10	13/13	100.00	160	0	100.0	25	
	M2wk	17	10/10	100.00	134	0	100.0	25	17
	M2wk	19	7/7	100.00	83	0	100.0	9	
	M4wk	24	2/2	100.00	22	0	100.0	4	2.5
	M4wk	23	1/1[d]	100.00	13	0	100.0	1	
	M8wk	27	*no peptides identified*						0.5
	M8wk	30	1/1[d]	100.00	13	0	100.0	1	
	M12wk	54	3/3[d]	100.00	40	0	100.0	3	3

Table 3.3: Least square regression of albumin and hemoglobin peptide abundance in post-feeding and post-molting experiment of *Triatoma protracta*.

	Term	Estimate	t Ratio	Prob > |t|
1. Post-feeding	Time	-51.33298	-7.97	**<0.0001***
	Molecule [Albumin, Hemoglobin]	-44.86667	-4.79	**<0.0001***
	Time x Molecule interaction	23.879202	3.71	**0.0010***
2. Post-molting	Time	-3.960947	-1.35	0.1938
	Molecule [Albumin, Hemoglobin]	-28.45455	-2.59	**0.0183***
	Time x Molecule interaction	1.3869822	0.47	0.6421

Supplementary Table I: Instances when the peptide identified was not known to match to the known blood meal source. All spectra that did not match the known blood meal source, *M. musculus*, were consistent with one of the two possible explanations: polymorphisms not previously reported for *M. musculus*, or peptides misidentified by SEQUEST. Manual examination of spectra was performed for these instances. For SEQUEST errors, the likely correct sequence is shown. HBA (hemoglobin alpha); HBB (hemoglobin beta), ALB (albumin), PF (post-feed), PM (post-molt)

Supplementary Table II: (A) SEQUEST identification of peptides. Here we show the hemoglobin peptides identified in both experiments, post-feeding and post-molting. (z) charge state; (#) acrylamidation on C; (*) oxidation on M; (^) Manually adjusted sequence to VNADEVGGEALGR-see Supplementary Table I and Supplementary Figure 1 for details. (B) Here we show the albumin peptides identified in both experiments, post-feeding and post-molting. (z) charge state; (#) acrylamidation on C; (*) oxidation on M; (^) Manually adjusted sequence to LGEYGFQNAILVR-see Supplementary Table I for details.

Supplementary Table III: Lab-reared *Triatoma protracta* were established into two groups representing the two experiments: (1) post-feeding or (2) post-molting. The two experimental groups were kept separately in equivalent containers. No non-experimental insects were housed in the experimental groups. The number of specimens present in the container is recorded at each time point, along with the number of specimens that died, and the number used for analysis.

Supplementary Table IV. Distribution of detection of hemoglobin and albumin by LC-MS/MS and DNA by SINE-PCR in post-feeding and post-molting *Triatoma protracta*. X/X is the number of positive blood meal vectors/total number of vectors at each time point.

Supplementary Figure 1: Critical ions in the fragmentation spectra are needed to distinguish VNADEVGGEALGR from NVADEVGGEALGR. A peptide which shows frequent mismatching by SEQUEST in mouse blood is VNADEVGGEALGR, as it is often identified with the first two amino acids inverted to NVADEVGGEALGR (Keller et al. 2017). We identify this peptide in our mouse-fed Triatominae and found that the critical b1, b2, and y12 ions that distinguish the first two amino acids are generally absent in the MS/MS spectra. When this is the case, the precursor peptide ions are often in low abundance (see Steen & Mann 2004 for details). However, when we perform targeted MS/MS analysis for the precursor mass of this peptide we regularly observe the y12 ion which allows one to appropriately determine the sequence as VNADEVGGEALGR. Indeed, we never observed fragment ions suggesting this peptide started with NV. This is strong evidence that VN and not NV is the proper solution for the N-terminal amino acids found in our samples. (A) The isotopic envelope for the precursor doubly-charged peptide with a theoretical monoisotopic mass of 643.820 *m/z*. The resulting MS/MS fragmentation spectrum (B) frequently includes some of the key diagnostic y12 fragment ion (C). When SEQUEST assigns the NV version of this peptide (D), key ions for determining the sequence of the first two amino acids is absent. The required y12 fragment ion of the NVADEVGGEALGR with a mass of 1172.591 is never identified (D,E), while the 1187.565 y12 fragment ion is identified when then peptide is targeted, allowing us to determine that VNADEVGGEALGR, which matches mouse, is indeed the correct sequence.

Supplementary Figure 2. Agarose gel of SINE-PCR product from (A) time post-feeding experiment and (B) time post-molting experiment.

3.8 References

1. Cantey PT, Stramer SL, Townsend RL, Kamel H, Ofafa K, Todd CW, et al. 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. 2012;52(9):1922-30.
2. Stimpert KK, Montgomery SP. Physician awareness of Chagas disease, USA. Emerg Infect Dis. 2010;16(5):871.
3. Edwards MS, Stimpert KK, Montgomery SP. Addressing the challenges of Chagas Disease: An emerging health concern in the United States. Infect Dis Clin Practice. 2017;25(3):118-25.
4. Lucero DE, Morrissey LA, Rizzo DM, Rodas A, Garnica R, Stevens L, et al. Ecohealth interventions limit Triatomine reinfestation following insecticide spraying in La Brea, Guatemala. Am J Trop Med Hyg. 2013;88(4):630-7.
5. Pellecer MJ, Dorn PL, Bustamante DM, Rodas A, Monroy MC. Vector blood meals are an early indicator of the effectiveness of the Ecohealth approach in halting Chagas transmission in Guatemala. Am J Trop Med Hyg. 2013;88(4):638-44.
6. Yoshioka K, Nakamura J, Pérez B, Tercero D, Pérez L, Tabaru Y. Effectiveness of large-scale Chagas disease vector control program in Nicaragua by residual insecticide spraying against *Triatoma dimidiata*. Am J Trop Med Hyg. 2015;93(6):1231-9.
7. Nakagawa J, Cordón-Rosales C, Juárez J, Itzep C, Nonami T. Impact of residual spraying on *Rhodnius prolixus* and *Triatoma dimidiata* in the department of Zacapa in Guatemala. Mem Ins Oswaldo Cruz. 2003;98(2):277-82.
8. Hashimoto K, Cordon-Rosales C, Trampe R, Kawabata M. Impact of single and multiple residual sprayings of pyrethroid insecticides against *Triatoma dimidiata* (Reduviiade; Triatominae), the principal vector of Chagas disease in Jutiapa, Guatemala. Am J Trop Med Hyg. 2006;75(2):226.
9. Gürtler RE, Yadon ZE. Eco-bio-social research on community-based approaches for Chagas disease vector control in Latin America. Trans R Soc Trop Med Hyg. 2015;109(2):91-8.
10. Waleckx E, Camara-Mejia J, Ramirez-Sierra MJ, Cruz-Chan V, Rosado-Vallado M, Vazquez-Narvaez S, et al. An innovative ecohealth intervention for Chagas disease vector control in Yucatan, Mexico. Trans R Soc Trop Med Hyg. 2015;109(2):143-9.
11. Vazquez-Prokopec GM, Spillmann C, Zaidenberg M, Kitron U, Gürtler RE. Cost-effectiveness of Chagas disease vector control strategies in northwestern Argentina. PLoS Negl Trop Dis. 2009;3(1):e363.
12. Gurevitz JM, Gaspe MS, Enriquez GF, Provecho YM, Kitron U, Gürtler RE. Intensified surveillance and insecticide-based control of the Chagas disease vector *Triatoma infestan*s in the Argentinean Chaco. PLoS Negl Trop Dis. 2013;7(4):e2158.

13. Henao-Martínez AF, Colborn K, Parra-Henao G. Overcoming research barriers in Chagas disease—designing effective implementation science. Parasitol Res. 2017;116(1):35-44.
14. Mesquita RD, Vionette-Amaral RJ, Lowenberger C, Rivera-Pomar R, Monteiro FA, Minx P, et al. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. Proc Natl Acad Sci USA. 2015;112(48):14936-41.
15. Mota J, Chacon JC, Gutiérrez-Cabrera AE, Sanchez-Cordero V, Wirtz RA, Ordonez R, et al. Identification of blood meal source and infection with *Trypanosoma cruzi* of Chagas disease vectors using a multiplex cytochrome b polymerase chain reaction assay. Vector-borne and Zoonotic Diseases. 2007;7(4):617-28.
16. Laskay ÜA, Breci L, Vilcins I-ME, Dietrich G, Barbour AG, Piesman J, et al. Survival of host blood proteins in *Ixodes scapularis* (Acari: Ixodidae) ticks: a time course study. J Med Ent. 2013;50(6):1282-90.
17. Pizarro JC, Lucero D, Stevens L. A method for the identification of guinea pig blood meal in the Chagas disease vector, *Triatoma infestans*. Kinetoplastid Biology and Disease. 2007;6:1.
18. Pinto J, Roellig DM, Gilman RH, Calderón M, Bartra C, Salazar R, et al. Temporal differences in blood meal detection from the midguts of *Triatoma infestans*. Rev Inst Med Trop Sao Paulo. 2012;54:83-8.
19. Kjos SA, Marcet PL, Yabsley MJ, Kitron U, Snowden KF, Logan KS, et al. Identification of bloodmeal sources and *Trypanosoma cruzi* infection in triatomine bugs (Hemiptera: Reduviidae) from residential settings in Texas, the United States. J Med Ent. 2013;50(5):1126-39.
20. Breniere SF, Pietrokovsky S, Gastelum EM, Bosseno MF, Soto MM, Ouaissi A, et al. Feeding patterns of *Triatoma longipennis* Usinger (Hemiptera, Reduviidae) in peridomestic habitats of a rural community in Jalisco State, Mexico. J Med Ent. 2004;41(6):1015-20.
21. Buitrago NLR, Bosseno MF, Waleckx E, Brémond P, Vidaurre P, Zoveda F, et al. Risk of transmission of *Trypanosoma cruzi* by wild *Triatoma infestans* (Hemiptera: Reduviidae) in Bolivia supported by the detection of human blood meals. Infection, Genetics and Evolution. 2013;19:141-4.
22. Peña VH, Fernández GJ, Gómez-Palacio AM, Mejía-Jaramillo AM, Cantillo O, Triana-Chávez O. High-resolution melting (HRM) of the cytochrome B gene: a powerful approach to identify blood-meal sources in Chagas disease Vectors. PLoS Negl Trop Dis. 2012;6(2):e1530.
23. Gonçalves TCM, Rocha DS, Cunha RA. Feeding patterns of *Triatoma vitticeps* in the State of Rio de Janeiro, Brazil. Rev de Saúde Pública. 2000;34(4):348-52.
24. Rabinovich JE, Kitron UD, Obed Y, Yoshioka M, Gottdenker N, Chaves LF. Ecological patterns of blood-feeding by kissing-bugs (Hemiptera: Reduviidae: Triatominae). Mem Inst Oswaldo Cruz. 2011;106(4):479-94.

25. Stevens L, Dorn PL, Hobson J, de la Rua NM, Lucero DE, Klotz JH, et al. Vector blood meals and Chagas disease transmission potential, United States. Emerg Inf Dis. 2012;18(4):646-9.
26. Georgieva AY, Gordon ERL, Weirauch C. Sylvatic host associations of Triatominae and implications for Chagas disease reservoirs: a review and new host records based on archival specimens. PeerJ. 2017;5:e3826.
27. Kieran TJ, Gottdenker NL, Varian CP, Saldaña A, Means N, Owens D, et al. Blood meal source characterization using Illumina Sequencing in the Chagas Disease vector *Rhodnius pallescens* (Hemiptera: Reduviidae) in Panamá. J Med Ent. 2017;54(6):1786-9.
28. Dumonteil E, Ramirez-Sierra M-J, Pérez-Carrillo S, Teh-Poot C, Herrera C, Gourbière S, et al. Detailed ecological associations of triatomines revealed by metabarcoding and next-generation sequencing: implications for triatomine behavior and *Trypanosoma cruzi* transmission cycles. Sci Rep. 2018;8(1):4140.
29. Keller JI, Ballif BA, Clair RMS, Vincent JJ, Monroy MC, Stevens L. Chagas disease vector blood meal sources identified by protein mass spectrometry. PLoS ONE. 2017;12(12):e0189647.
30. Lucero DE, Ribera W, Pizarro JC, Plaza C, Gordon LW, Pena R, Jr., et al. Sources of Blood Meals of Sylvatic *Triatoma guasayana* near Zurima, Bolivia, Assayed with qPCR and 12S Cloning. PLoS Negl Trop Dis. 2014;8(12):e3365.
31. Greenwalt DE, Goreva YS, Siljestrom SM, Rose T, Harbach RE. Hemoglobin-derived porphyrins preserved in a Middle Eocene blood-engorged mosquito. Proc Natl Acad Sci USA. 2013;110(46):18496-500.
32. Laskay ÜA, Burg J, Kaleta EJ, Vilcins IM, Telford Iii SR, Barbour AG, et al. Development of a host blood meal database: de novo sequencing of hemoglobin from nine small mammals using mass spectrometry. Biol Chem. 2012;393(3):195-201.
33. Lehane MJ. The biology of blood-sucking in insects: Cambridge University Press; 2005.
34. Belles X. Origin and evolution of insect metamorphosis. Encyclopedia of Life Sciences (ELS). Chichester, UK: John Wiley & Sons, Ltd; 2011.
35. National Research Council. Guide for the care and use of laboratory animals: National Academies Press; 2010.
36. Feliciangeli MD, Rabinovich J, Fernandez E. Resistencia al ayuno em Triatominos (Hemiptera, Reduviidae) venezolanos. Rev Inst Med Trop Sao Paulo. 1980;22(2).
37. Ballif BA, Carey GR, Sunyaev SR, Gygi SP. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. J Prot Res. 2008;7(1):311-8.
38. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2009;37(Database issue):D26-31.
39. Walker JA, Hughes DA, Hedges DJ, Anders BA, Laborde ME, Shewale J, et al. Quantitative PCR for DNA identification based on genome-specific interspersed repetitive elements. Genomics. 2004;83(3):518-27.

40. Pizarro JC, Stevens L. A new method for forensic DNA analysis of the blood meal in Chagas disease vectors demonstrated using *Triatoma infestans* from Chuquisaca, Bolivia. PLoS ONE. 2008;3(10):e3585.
41. Gürtler R, Cecere M, Lauricella M, Cardinal M, Kitron U, Cohen J. Domestic dogs and cats as sources of *Trypanosoma cruzi* infection in rural northwestern Argentina. Parasitol. 2007;134(01):69-82.
42. Önder Ö, Shao W, Kemps BD, Lam H, Brisson D. Identifying sources of tick blood meals using unidentified tandem mass spectral libraries. Nature Communications. 2013;4:1746.
43. Niare S, Berenger J-M, Dieme C, Doumbo O, Raoult D, Parola P, et al. Identification of blood meal sources in the main African malaria mosquito vector by MALDI-TOF MS. Malaria Journal. 2016;15(1):1.
44. Noriega FG. Autogeny in three species of Triatominae: *Rhodnius prolixus, Triatoma rubrovaria,* and *Triatoma infestans* (Hemiptera: Reduviidae). J Med Ent. 1992;29(2):273-7.
45. Nattero J, Leonhard G, Rodríguez CS, Crocco L. Influence of the quality and quantity of blood ingested on reproductive parameters and life-span in *Triatoma infestans* (Klug). Acta Trop. 2011;119(2):183-7.
46. Zeledón R, Guardia VM, Zuñiga A, Swartzwelder JC. Biology and ethology of *Triatoma Dimidiata* (Latreille, 1811) II. Life span of adults and fecundity and fertility of females. J Med Ent. 1970;7(4):462-9.
47. Gürtler RE, Cecere MC, Vazquez DP, Chuit R, Cohen JE. Host-Feeding Patterns of Domiciliary *Triatoma infestans* (Hemiptera: Reduviidae) in Northwest Argentina: Seasonal and Instar Variation. J Med Ent. 1996;33(1):15-26.
48. Gürtler RE, Cardinal M. Reservoir host competence and the role of domestic and commensal hosts in the transmission of *Trypanosoma cruzi*. Acta Trop. 2015;151:32-50.
49. Hardison RC. Evolution of hemoglobin and its genes. Cold Spring Harbor Perspectives in Medicine. 2012;2(12):a011627.
50. Luz C, Fargues J, Romana C. Influence of starvation and blood meal-induced moult on the susceptibility of nymphs of *Rhodnius prolixus Stål* (Hem., Triatominae) to *Beauveria bassiana* (Bals.) Vuill. infection. J Appl Ent. 2003;127(3):153-6.
51. Martínez-Ibarra JA, Grant-Guillén Y, Nogueda-Torres B, Trujillo-Contreras F. Influence of the blood meal source on the biology of *Meccus longipennis* (Hemiptera: Reduviidae) under laboratory conditions. J Am Mosq Control Assoc. 2004;20(3):328-30.
52. Post R, Flook P, Millest A. Methods for the preservation of insects for DNA studies. Biochemical Systematics and Ecology. 1993;21(1):85-92.
53. Padrón G, Domont GB. Two decades of proteomics in Latin America: a personal view. Journal of Proteomics. 2014;107:83-92.

CHAPTER 4: PROTEIN MASS SPECTROMETRY OUTPERFORMS DNA-BASED METHODS FOR DETECTION AND QUANTIFICATION OF BLOOD MEAL SOURCES TO ENHANCE CHAGAS VECTOR ECOHEALTH MANAGEMENT

Judith I. Keller[1], Raquel Lima-Cordón [2,1], M. Carlota Monroy [2,1], Anna M. Schmoker[1], Fan Zhang[1], Bryan A. Ballif [1,*], and Lori Stevens [1,*]

[1] Department of Biology, University of Vermont, Burlington, Vermont, United States of America

[2] Laboratorio de Entomología Aplicada y Parasitología, Escuela de Biología, Facultad de Ciencias Químicas y Farmacia, Universidad de San Carlos de Guatemala, Edificio T-10 Ciudad Universitaria Zona 12, Ciudad de Guatemala, Guatemala

*Co-corresponding authors:

Email: lori.stevens@uvm.edu (LS), bballif@uvm.edu (BAB)

4.1 Abstract

Chagas disease, a neglected tropical disease endemic in Latin America, is caused by the protozoan parasite *Trypanosoma cruzi* and is responsible for significant health impacts, especially in rural communities. The parasite is transmitted by insect vectors in the Triatominae family and due to lack of vaccines and limited treatment options, vector control is the main way of controlling the disease. Knowing on what vectors are feeding on directly enhances our understanding of the ecology and biology of the different vector species and can potentially aid in engaging communities in active disease control, a concept known as Ecohealth management. We evaluated blood meals in field-caught insect vectors previously evaluated for blood meals via DNA analysis as part of a larger collaborative project in Guatemala, Central America. In addition to identifying blood meals in 100 % of all samples using a mass spectrometry-based approach (liquid chromatography tandem mass spectrometry - LC-MS/MS), we identified blood meals in 53 % of insect vectors that had no evidence of a recent blood meal by DNA-PCR. As single vectors often feed on multiple sources, we validated our ability to decipher multiple blood meals from an individual vector and showed the ability to quantify a blood meal using synthetic AQUA (Absolute QUAntification) peptides, a first step in further exploration of species-specific blood meal composition. Furthermore, we show that a lower resolution mass spectrometer is able to identify blood meals from taxa correctly, an important and strong attribute of our LC-MS/MS-based method, opening the door to using proteomics in countries where Chagas disease is endemic and resources are limited.

4.2 Introduction

Chagas disease is a neglected tropical disease of great importance in Latin America, affecting millions of people (World Health Organization 2016, World Health Organization and UNICEF 2017), including those in rural villages of Guatemala, Central America. While the acute phase of this disease usually only causes mild symptoms of malaise, chronic disease complications, with cardiomyopathy being the most prevalent, have significant impacts on health and well-being in local communities. Chagas disease is the number one cause for heart problems in these parts of the world (World Health Organization 2016). As the *Trypansosoma cruzi* parasite, the etiological agent of Chagas disease, is transmitted mainly by an insect vector, this disease involving multiple players can be challenging to control: the Triatominae family of vectors, colloquially known as 'kissing bugs', contains over 150 different species (Justi and Galvão 2017), and as far as is known all with the potential of carrying the Chagas parasite, and with the ability to infect any mammal.

In locations where local government resources are often minimal, it can be challenging to implement large-scale disease management strategies. Although vector management through large-scale insecticide spraying has occurred in the past and is still happening today, the effectiveness of this single strategy in controlling the vector seem to be limited to introduced species outside their native ranges (e.g., *Rhodnius prolixus*, *Triatoma infestans*) (Nakagawa et al. 2003, Cecere et al. 2006). Native vector management, however, such as that of *Triatoma dimidiata* which is abundant in Central America and especially Guatemala, faces different challenges, such a rapid

recolonization of houses from sylvatic habitats following insecticide spraying (Hashimoto et al. 2006).

In recent years, the concept of implementation science and using Ecohealth-based approaches has been recognized as a sustainable method of Chagas disease control. Involving affected communities in Central and South America, (Pellecer et al. 2013, Waleckx et al. 2015), and using education coupled with scientific data, has had an impact at various scales in controlling disease burdens in rural areas. In addition, blood meal data and vector prevalence of various hosts fosters understanding of local vector ecology and behavior, and ultimately allows communities to make sustainable decisions for not only managing Chagas disease, but also decreasing helminth infections and increasing nutrition.

Various methods to detect vector feeding profiles have been applied to the triatomine vectors of Chagas disease. While protein-based methods were commonly used in the earliest studies of blood meal detection in vectors (i.e., precipitin and antisera tests, recently evaluated in (Rabinovich et al. 2011)), DNA-based detection methods have come to the forefront. Many of these methods are based on the relatively abundant mitochondrial DNA or repetitive nuclear DNA regions (Kent 2009). The polymerase chain reaction (PCR), sometimes coupled with DNA sequencing, using species-specific or general target primers (i.e., the mitochondrial 12 S ribosomal gene) is common (Pizarro et al. 2007, Pizarro and Stevens 2008, Lucero et al. 2013, Klotz et al. 2014, Lucero et al. 2014, Lima-Cordón et al. 2018). Indeed, we have previously evaluated if various locations in Central America differ in their role of sylvatic,

synanthropic, and domestic animals in the local Chagas transmission cycles as part of an Ecohealth management approach in these areas using taxa specific primers to detect blood meal prevalence (Lima-Cordón et al. 2018).

However, DNA-based blood meal detection methods often come with the challenge of needing uncontaminated, high-quality DNA, often in large quantities from recently-fed vectors (Stevens et al. 2012). Proteomics studies are becoming more frequent in determining blood meal sources from various arthropod disease vectors (Laskay et al. 2013, Önder et al. 2013, Niare et al. 2016), and we have previously shown the application of a liquid chromatography tandem mass spectrometry (LC-MS/MS)-based approach to determining blood meal sources in insect vectors (Keller et al. 2017, Keller et al. 2018). Proteomics resources are available in most Chagas endemic areas (Padrón and Domont 2014) and previous cost analysis demonstrates an LC-MS/MS-based approach to blood meal identification is very realistic for Latin America (Önder et al. 2013, Keller et al. 2017). In addition, LC-MS/MS targeting hemoglobin protein peptides found in the blood meal have allowed us to identify blood meals over long temporal scales and even post-molting to the limits of our experimental replicates (4 and 12 weeks, respectively) (Keller et al. 2018). Hemoglobin peptide signatures lasted longer than DNA (Keller et al. 2018) when compared short interspersed nuclear element (SINE)-DNA PCR and LC-MS/MS always detected blood meals in instances where no blood meal was evident using DNA-based methods (Keller et al. 2018 – SINE PCR, Keller et al. 2017 – 12 S PCR).

With 12 S mitochondrial sequencing, multiple blood meals are often detected in Triatominae vectors (Georgieva et al. 2017). Detecting multiple blood meals with 12 S sequencing, however, proves challenging and an additional cloning step is required when multiple blood meals are present (Lucero et al. 2014, Waleckx et al. 2014). Although genomics-based approaches are emerging (Collini et al. 2015, Logue et al. 2016, Kieran et al. 2017, Dumonteil et al. 2018, Orantes et al. 2018), cost, sample processing time, the window of time to detect multiple blood meals, and contamination are challenges that need to be overcome. Another strong advantage of mass spectrometry-based methods is the ability of a single LC-MS/MS run to identify all/multiple blood meals at the same time, rather than with single PCR for each taxa, or cloning and sequencing reactions, which is something we address in this study.

In this study we develop the ability to (1) detect and identify single and multiple blood meal sources in field-collected *Triatoma dimidiata* and *Triatoma nitida* from Jutiapa, Guatemala that are part of a large-scale Ecohealth intervention project using a hemoglobin peptide-based LC-MS/MS approach. In addition to identifying blood meal sources, we (2) develop and validate our multiple blood meal detection pipeline and quantify blood meals using synthetic Absolute Quantification (AQUA) peptides as controls. Lastly, as the blood meal profiles of the field-collected *T. dimidiata* and *T. nitida* specimens used in this study have been previously investigated using taxa-specific primers by PCR, we (3) compare the detection efficiency of LC-MS/MS for samples where DNA-based methods both detected and did not detect a blood meal.

4.3 Methods

4.3.1 Ethics Statement

All ethics of previously investigated samples are contained within Lima-Cordón et al. 2018. In short, ethical clearance was granted from the Ministry of Health in Guatemala and the PanAmerican Health Organization. In addition, a single mouse blood sample used for this study was obtained in accordance with an IACUC-approved protocol encouraging "tissue sharing" of post-mortem tissue. All procedures were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Vermont in accordance with the requirements of the Office of Laboratory Animal Welfare (IACUC protocol 12-045). The mouse blood used in this study was previously described in Keller et al. 2018, and mouse blood was stored in 100mM ethylenediamine tetraacetic acid (EDTA) (Sigma-Aldrich, Saint Louis, Missouri, USA). The mouse blood sample was processed as previously described (Keller et al. 2018) and run on the mass spectrometer as described below.

4.3.2 Insect vector collection and storage

Field-collected *Triatoma dimidiata* and *Triatoma nitida* specimens were collected and analyzed previously as part of a larger collaborative project "Ecohealth interventions for the prevention of Chagas Disease in Central America" (https://www.idrc.ca/en/project/ecohealth-interventions-chagas-disease-prevention-central-america-0) (for pre-Ecohealth intervention baseline survey data see (Bustamante-Zamora et al. 2015, Lima-Cordón et al. 2018). For 45/50 specimens used in this study, corresponding base-line survey data were available in regards to animal

practices in the household (Table 4.1) (Bustamante-Zamora et al. 2015, Lima-Cordón et al. 2018).

A subset of specimens used in Lima-Cordón et al (2018) was used in this study and analyzed for blood meal sources. These specimens were chosen for comparison because PCR is able to detect a blood meal for shorter time post-feeding compared to LC-MS/MS (Keller et al. 2018). Many of the specimens had no evidence of a recent blood meal detected with DNA, thus we were able to enhance the Ecohealth information provided by these samples. Specimens that were positive for a blood meal by DNA were chosen randomly, while we chose some engorged and some starved specimens for samples that were negative for a blood meal with DNA (Table 4.1). A few specimens were not reported on in Lima-Cordón et al. (2018) for various reasons (i.e. different vector species, no insect age recorded, etc.) were included in this study since blood meal species identification by DNA data were readily available. To compare LC-MS/MS with DNA, the last three segments of the abdomen were used for blood meal analysis, while for younger life stages, slightly more cranial abdomen pieces were used for LC-MS/MS analysis if a vector was exceptionally small.

The fifty insects examined were from houses and peridomestic areas and transported to the Laboratory of Applied Entomology and Parasitology (LENAP), Universidad the San Carlos de Guatemala, Ciudad de Guatemala, Guatemala as previously described (Lima-Cordón et al. 2018) in August-October 2011, stored in 95 % ethanol and 5 % glycerol at room temperature, and transported to the University of Vermont in June 2016 and March 2017. Specimens were subsequently stored at 4 °C

for up to 9 months, of note, most of the ethanol/glycerol mixture had evaporated by the time the specimens arrived at the University of Vermont, and were processed for mass spectrometry analysis within 1-9 months.

4.3.3 Blood meal detection by DNA

In order to compare LC-MS/MS and PCR of DNA methods for blood meal source detection, we examined 43 *T. dimidiata* and 7 *T. nitida* specimens previously examined by DNA (Lima-Cordón et al. 2018). We briefly summarize the previously reported DNA methods here. PCR products from reactions using taxa specific primers for human, dog, bird, mouse, rat, opossum, and pig were visualized with agarose gel electrophoresis and transillumination of appropriately sized bands (Lima-Cordón et al. 2018). Most of the taxa-specific PCR assays targeted highly repetitive short interspersed nuclear element (SINE) regions (Supplementary Table 4.1). Of the 568 specimens previously examined 30 % (n=170) were found to have evidence of a recent blood meal by PCR and of these, 7 % (n=12) had evidence of two blood meal sources per specimen (Lima-Cordón et al. 2018). In the study reported here, we further examined 23 samples with a previous blood meal (4 samples had 2 blood meal sources detected) and 27 samples for which no blood meal was detected by PCR. Two of the samples with no previous blood meal detected had also been re-examined by PCR and sequencing targeting the mitochondrial 12 S genes, one identifying human and one identifying cow as the blood meal source.

4.3.4 Hemoglobin peptide detection by LC-MS/MS

We extracted protein from the fifty specimens as previously described (Keller et al. 2018). Briefly, abdominal insect tissue was ground in 95 °C denaturing buffer, boiled and spun down. Denatured proteins were separated by 15 % SDS-PAGE and cut from gel regions surrounding the molecular weight of hemoglobin (~16kDa). These regions were subjected to in-gel digestions with trypsin (Keller et al. 2017, Keller et al. 2018). Additionally in this study, samples were reconstituted with 20 µL 2.5 % acetonitrile/2.5% formic acid rather than 9 µL for LC-MS/MS analysis to avoid overloading the column. LC-MS/MS uses tandem mass spectrometry to assay masses from fragmented peptides. Our assays used a linear ion trap-orbitrap (LTQ-Orbitrap; Thermo Electron, Waltham, Massachusetts, USA) to identify peptides where all spectra were collected in the orbitrap (Ballif et al. 2008, Keller et al. 2018). In order to examine the ability of lower resolution mass spectrometry instruments for blood meal identification, three samples (JKop16_TPG097, JKop17_TPG100, mouse blood 1996) were subjected to an additional run with all spectra collected in the LTQ.

As described previously, we searched mass spectra using the SEQUEST algorithm (Thermo Electron V26.12) against a custom forward and reverse concatenated database containing vertebrate hemoglobin sequences (17 Jan 2018, 18,000+ entries) extracted from GenBank using ‘hemoglobin’ as a keyword in any curated field and ‘vertebrate’ as a filter (Keller et al. 2017, Keller et al. 2018). Peptide filter parameters in SEQUEST were also as described previously, except for the three

samples run in the lower resolution LTQ instrument where the measured precursor masses were changed from ± 5 ppm to 2 Da.

Some trypsin digested hemoglobin peptides are variable within species, while others are invariant within species but variable between species. The unique peptide sequences identified for each specimen were indexed with the corresponding taxonomic information extracted from GenBank as previously described using Pyteomics python tools and library (Keller et al. 2017). Prior to indexing, three aspects of the database were further curated. First, because we include peptides with hemoglobin in any field from GenBank, not just curated entries with e.g., gene = hemoglobin, we further curated our database through BLAST searching each unique peptide identified and confirming it was a hemoglobin sequence. Non-hemoglobin peptides were not considered for downstream analysis. Second, taxonomic categories vary among vertebrates (e.g., it is debated whether the monophyletic clades artiodactyla and cetacea are more likely comprised of the non-monophyletic clade cetartiodactyla), thus we added taxonomic information where blanks existed for downstream analysis (Supplementary Table 4.2). Finally, a few peptides (LLGNVXVCVLAHHFGK, SAVTAXWGK, (K)VLNSFSDGXK) included an "X" amino acid, signifying either leucine or isoleucine in the sequence. We individually adjusted these "X" amino acids after determining they were leucine (L) through a BLAST search.

We indexed unique peptides against the associated taxonomy using JMP, Version 14 (SAS Institute Inc., Cary, NC, 1989-2018). Below we describe an updated,

more streamlined pipeline that addresses the determination of multiple blood meals identified in this study (Figure 4.1).

4.3.5 Determining blood meals from hemoglobin peptide data

In our previous LC-MS/MS studies we only had evidence for a single blood meal source in each specimen examined (Keller et al. 2017, Keller et al. 2018). However, the DNA-based study of the specimens examined in this study had reported evidence of two blood meal sources in four of the specimens (Lima-Cordón et al. 2018). Therefore, we describe how we expanded the data analysis pipeline to identify multiple blood meal sources in a single specimen.

Using our indexed list of unique peptides from each specimen and associated taxonomic data, we examined the number of classes, orders, families, genera, and species represented in each sample and found taxonomic order to be the most informative for the first step in determining multiple blood meals and made a histogram with the number of unique peptides in each of the orders represented in a sample. If sets of peptides for any order were an exact subset of an order with more unique peptides, that order was no longer considered a blood meal group. Similarly, if a set of peptides was a subset of two or more orders with more unique peptides, that order was no longer considered a unique blood meal group (Figure 4.1 E.1).

First, we describe how we identify major blood meals. For a blood meal source to be considered from a particular order, it has to be either the most supported order (maximum number of unique peptides N_{max}) or contain more than $N_{50\%max}$ unique peptides (Figure 4.1 E). By only considering the orders with 50 % of the number of

peptides of the most supported order, in addition to only considering blood meal orders where unique peptides exist for an order, we maintain high stringency and minimize the false discovery rates, or minimize falsely inferring an additional blood meal. We refer to this as the 'top 50 %' approach. See Figure 4.1 for more detail.

Our analysis becomes problematic when there are unique taxonomic orders and their associated peptides that do not fall within the determined major blood meal orders. For example, the peptide VADALTTAVSHIDDLPGALSALSDLHAYK found in seal has been determined to be an unreported polymorphism for dog (Keller et al. 2017). Another problem is misidentification of the peptide sequence by the SEQUEST program. For example, the peptide VNADEVGGEALGR is often identified by SEQUEST as NVADEVGGEALGR (see Keller et al. 2018 for details). Experience examining the data show that there are usually very few peptides represented for these mistakenly suggested orders and these do not hamper correctly identifying a blood meal sources. In addition to the two aforementioned scenarios, peptides could also represent rare, or of low abundance "minor" blood meals. Thus, taxa not falling under the top-50% approach may fall into one of 3 categories: (1) minor (rare/low-abundance) blood meals; (2) previously unknown polymorphisms; (3) mis-identifications by SEQUEST.

4.3.6 AQUA peptide analysis

We have previously demonstrated our ability to quantitatively profile Triatominae blood meal sources using stable isotope-labeled hemoglobin peptides, also known as AQUA (Absolute QUAntification) (Keller et al. 2017). We extrapolate here on this concept and use AQUA peptides to (1) validate our ability to detect multiple

blood meals, and to (2) generally quantify total amount of hemoglobin, using absolute amounts of highly conserved hemoglobin sequences as a proxy for the total amount of hemoglobin present in a sample.

A subset of 11 samples were run with AQUA peptides and LC-MS/MS methods were as previously described (above, Keller et al. 2017 and Keller et al. 2018), except that for the AQUA peptides a top-2 approach was used, where the top two most abundant ions were collected, followed by a targeted approach where we target the dominant charge state of each native peptide for which we have spiked in a corresponding synthetic AQUA peptide (Table 4.2). All spectra were collected in the orbitrap. Concentration (in pmol/µl) were known of each AQUA peptide from the manufacturer, and a 10-AQUA peptide mix was optimized to run with insect vector samples. Due to overlapping species matches, only 5 AQUA peptides were included in the analysis presented here.

Validation of multiple blood meals: From the previously published DNA-based blood meal identification of samples used in this study (Lima-Cordón et al. 2018), we synthesized AQUA peptides based on alpha and beta hemoglobin sequences that help distinguish between different target blood meal sources (Table 4.2) at Cell Signaling Technology (Danvers, Massachusetts, USA). Since the vast majority of blood meals in Lima et al. 2018 were from human, chicken, and dog, we constructed AQUA peptides for each of these species. Each AQUA peptide contained one heavy-labeled amino acid (Table 4.2). We also included a more general bird AQUA peptide that is common throughout 15 avian orders. As there are not necessarily unique peptide sequences

available for single species, most AQUA peptides were chosen with the goal of clear distinction between the three target species (human, chicken, dog). Therefore, for this particular experimental population, while an AQUA may match a target species, it may also match a small number of other species with overlapping sequences. Presence/absence data of native vs. AQUA MS1 peaks was collected for species-specific AQUA peptides, which served to validate our blood meal identification pipeline.

General quantification of a blood meal: We also constructed AQUA peptides of highly conserved regions of hemoglobin, where peptide sequences corresponded to many different taxa (Table 4.2). This offers a straightforward way to quantify the absolute amounts of a highly conserved peptide, a proxy for the total amount of hemoglobin present in a sample and overall quantification of the blood meal. We chose this sequence based on the low decay rate over time in ticks as shown in (Laskay et al. 2013) in ticks. Quantification of a blood meal does not necessarily help in determining when an insect vector fed, as varying amounts of blood can be taken up depending on the length of the blood meal and the organism from which the blood is drawn. However, this is the first step toward using AQUA quantification for accurately profiling vector blood meals and elucidating blood meals not present in GenBank as suggested in (Keller et al. 2017).

In order to quantify blood meals, absolute amounts (in fmol) of synthetic and native peptides were measured and relative abundances calculated based on the heavy-to-light ratio of synthetic to native peptide comparing monoisotopic peak intensities for

highly conserved regions of hemoglobin, or general peptides. Since we and others (Lundgren et al. 2010, Keller et al. 2018), have used spectral counts, or the number of peptides identified in a sample as a proxy for protein quantification, we compared the fmol amount of general peptides identified (a representation for total amount of hemoglobin identified in a sample) to the spectral count of all peptides identified in a sample. We have previously instated stringent peptide filters that serve for very accurate blood meal species identification, however, we recognized that some of our XCorr filters were too stringent for some of the best AQUA fragmentation spectra. Therefore, we relaxed the constraints on our peptide filters as follows in order to get a more accurate representation of all hemoglobin peptides in a sample. These lower, but still stringent filters are as follows: (1) XCorr values greater than or equal to 1.8 (z=1), 2 (z=2), 2.2 (z=3), 2.4 (z=4), 2.5 (z=5). Empirically in large data sets, reducing XCorr values by ~0.2 per charge state, reduces false-discovery rates equally across all charge states.

4.4 Results

In this study we showed that hemoglobin peptide-based LC-MS/MS could detect and identify Chagas vector blood meal sources of field-collected Triatomine vectors from Guatemala and outperformed classical DNA-based PCR. First, we developed an updated pipeline for identifying single and multiple blood meals, which we then validated using synthetic AQUA peptides as positive controls. Once we validated our pipeline, we identified blood meals across all 50 samples using LC-

MS/MS, and directly compared between LC-MS/MS and DNA-identified vector blood meals. Lastly, we quantified blood meals using general AQUA peptides, and showed a direct correlation between hemoglobin peptide amount (fmol) and peptide spectral count, the number of hemoglobin peptides identified in a sample.

4.4.1 AQUA peptide ms1 spectra confirm presence and absence of multiple blood meals

To evaluate our ability to correctly identify multiple vector blood meals using our developed sample pipeline, a subset of 11 samples was re-run with AQUA peptides as controls. Based on the previously published DNA data (Lima-Cordón et al. 2018), the most common blood meal sources were human, bird (i.e. chicken), and dog. As we were expecting a similar distribution of blood meal sources in the 50 samples we evaluated from Lima-Cordón et al. 2018, we spiked AQUA peptides into 11 samples corresponding to the aforementioned most commonly identified blood meals (Figure 4.2, Table 4.2, Table 4.3).

With the exception of a few blood meal species (Table 4.3 *), the spiked-in control AQUA peptides allowed us to verify our blood meal species pipeline. As expected, for samples identifying as chicken/turkey (JKop53_TPG017, JKop27_TPG630, JKop22_TPG168, JKop21_TPG158), native peptide peaks were present for the chicken/turkey AQUA peptide as well as the general bird AQUA peptide. For sample JKop23_TPG186, the general bird peptide was present while the chicken/turkey peptide was absent even though the blood meal was identified as chicken/turkey. However, on closer examination, this particular sample contained a

slight variant or polymorphism in the AQUA peptide FLSAVSAVLAEK, where a serine was substituted at position 11 of this particular peptide. This would explain why no native peak was detected for chicken/turkey in this sample. JKop19_TPG104 and JKop09_TPG075 did not have a native human peptide identified, however in both cases there were significantly less human peptides identified than the primary blood meal, dog, which could lead to the human peptide being of too low abundance to be detected here. In samples where blood meals were identified that did not have a corresponding AQUA peptide, no peaks were appropriately detected for the 3 target species. For sample JKop30_TPG045 which identifies as a duck, the general avian native peptide was present while the chicken/turkey peptide was absent.

4.4.2 LC-MS/MS and DNA-identified vector blood meals

After being confident in our blood meal detection pipeline, we identified blood meal species in 50 Triatomine vectors (Figure 4.3, Supplementary Table 4.3, Supplementary Table 4.4).

Generally, DNA and LC-MS/MS-identified blood meal sources agreed on species, with the major exception being the pig primer (Figure 4.3, Supplementary Table 4.4). A few other discrepancies included JKop006_TPG038 and JKop27_TPG630 where human was detected with DNA but not with LC-MS/MS. DNA-based identification detected opossum and cow in sample JKop22_TPG168 while LC-MS/MS identified chicken/turkey (Figure 4.3, Supplementary Table 4.3, Supplementary Table 4.4). Overall, however, LC-MS/MS detected significantly more

blood meals than did DNA (100 % vs. 46 %, n=50) (Figure 4.3). Seventy percent of the specimens only able to be identified by LC-MS/MS contained human blood meals.

Not only did LC-MS/MS identify more blood meals overall (Figure 4.3), but also identified more blood meals per individual vector (Figure 4.4). Multiple major blood meals were detected in 46 % of specimens (n=50), while 28 % had 2 major blood meals and 18 % had 3 major blood meals. In comparison, only 4 samples contained 2 blood meals with DNA (8 %).

For some samples, top primate species were related species in the Hominidae family. For example, JKop045_TPG603 had equal support for *Pan paniscus* (bonobo) and *Homo sapiens* (human). As bonobos are native to Africa, the blood meal was determined to have come from a human based on the ecology of the species (Supplementary Table 4.3). The same principle applies for JKop33_TPG244 where 7 different primates were equally identified, yet the only ecologically relevant and local species is human. Significantly lower peptide support was frequently seen for some taxa of rodents that were not able to be identified to species (Rodentia/Cricetidae or Sciuridae) (Supplementary Table 4.3).

4.4.3 Survey data

For 45/50 samples, pre-intervention survey data on bio-socio-ecological factors existed in relation to animal practices within a household, totaling to 32 houses (Figure 4.5). Five houses had more than one vector collected for blood meal analysis (Table 4.1, a-e), while for all other houses only one vector was analyzed. Households where vectors were collected owned 0-12 dogs (mean 5.8, median 6, mode 4), 0-25 birds

(mean 2.5, median 2, mode 2), 0-2 pigs (mean 0.07, median, 0, mode 0), and 0-1 cows/beasts (mean 0.16, median 0, mode 0) (Table 4.1, Figure 4.5). All houses were inhabited (potential human blood source), while evidence for mice/rats/rodents was found in all but 2 houses (94 %) (Figure 4.5). All houses, except 2 owned birds (94 %), all but 6 owned dogs (81 %), while only 16 % owned pigs or other livestock.

4.4.4 Lower resolution mass spectrometer can identify multiple blood meals

Resources in Latin American countries where Chagas disease is endemic can be limited, and although proteomics equipment is present in many countries (Padrón and Domont 2014), peptide confidence and resolution of the instrument in terms of accurate mass measurements can influence the adaptability of our LC-MS/MS technique. Therefore, we wanted to evaluate the ability to detect the blood meal source(s) from (1) a control blood source (mouse – 1996); (2) a single blood meal (JKop016_TPG097) and (3) a multiple blood meal (JKop17_TPG100). We ran each sample once on the orbitrap and filtered peptides as previously described in this study and (Keller et al. 2017, Keller et al. 2018), and also on the lower mass resolution LTQ instrument, filtering peptides as with the orbitrap, but lifting the ppm restriction. Overall, the orbitrap identified samples to species level with high accuracy, while the LTQ identified blood meal sources well to at least order with high confidence, and some to family and species level (Supplementary Table 4.5).

4.4.5 Stringently filtered peptides give most accurate blood meal species identification

Two samples identified with LC-MS/MS (JKop35_TPG288 and JKop52_TPG814) had a blood meal detected and hemoglobin peptides were present, yet the blood meal was not able to be resolved to blood source species and the spectral count, or number of peptides identified, was extremely low. DNA did not detect a blood meal in either of two samples. Although our stringent peptide filters give the most accurate species identification, we evaluated these 2 samples in conjunction with 5 additional samples with lower peptide filtering stringency. Slightly lowering the peptide filter stringency allowed for species-level blood meal identification in both samples (Supplementary Table 4.6).

However, the slightly lower stringency peptide filters led to more orders being considered for major blood meals, sometimes incorrectly (Supplementary Table 4.6). This was most likely to due to a higher number of unanticipated polymorphisms. In addition, more polymorphic peptides per sample led to species-specific support to be of lower percentage in terms of unique peptides identified than with high stringency peptide filters. For multiple blood meals, highly stringently filtered peptide data gives more accurate species identification.

4.4.6 Blood meal quantification correlates to spectral count

Quantification of blood meals could lead to elucidating blood meal sources not currently in our underlying database, GenBank, and the experiment presented here is the first step in this aim by quantifying general hemoglobin peptides.

When evaluating our general AQUA peptides LLIVPWTQR/ LLVVYPWTQR for quantification, we realized the XCorr restrictions placed with our stringent peptide filters were not allowing an accurate representation of all hemoglobin peptides identified in a sample. We therefore used our lower stringency filters to get an accurate representation of hemoglobin peptides identified in a sample. Overall, the fmol amount of general hemoglobin peptides LLIVPWTQR/ LLVVYPWTQR had a positive correlation with the number of peptides identified in a sample (Figure 4.6). Only non-saturated data points were included, as spectral counts can be limited by the capacity of the ion trap used due to the total amount of protein (i.e., including insect proteins, etc.) present in a sample.

4.5 Discussion

Due to the limited availability of treatments for Chagas disease, coupled with the prevalence of Chagas disease in rural communities of Latin America, achievable, alternative disease management methods are highly sought after. Implementation science techniques such as Ecohealth approaches have proven to decrease disease burdens at local scales (Pellecer et al. 2013, Waleckx et al. 2015) and can serve to the overall betterment of affected communities. One step in utilizing sustainable Ecohealth management is identifying on what local insect vectors are feeding on. Knowing the vector blood source is crucial to understanding local transmission cycles, identifying clinically relevant host reservoirs, understanding the ecology and behavior of insect vectors, all leading to inform vector control decisions.

Herein we describe an efficient and straight-forward method for identifying multiple vector blood meals using protein mass spectrometry-based LC-MS/MS targeting hemoglobin molecules, which outperformed DNA-based methods in more than half of all samples.

In this study we detected blood meals in 100 % of samples using LC-MS/MS. This is in stark comparison to PCR-based methods which only detected blood meals in roughly half of the subset of specimen used in this study. Overall, blood meal species were detected in 28 % of specimens in Guatemala in context of the entire study (Lima-Cordón et al. 2018) (Figure 4.3), a percentage similar to other studies where overall blood meal detection appears to be low. Our LC-MS/MS-identified blood meals enhanced the ability to make informed vector management decisions. For example, Lima-Cordón et al. 2018 detected that vectors in Guatemala were feeding on humans, birds, dogs, and rodents in 17 %, 31 %, 45 %, and 3% respectively, while from the subset of 50 specimens analyzed here, 56 % of vectors (n=28) had taken a blood meal from a human, 42 % from chicken/turkey (n=21), 34 % from dog (n=17), and 18 % from a rodent (n=9). Overall, LC-MS/MS identified significantly more human blood meals. While with DNA-based detection methods contamination can occur, which was possibly the case for the 12 S blood meal identified sample (JKop06_TPG038 and JKop27_TPG630), contamination is extremely rare when using LC-MS/MS due to targeting hemoglobin molecules.

Bio-socio-ecological factors can influence domiciliary infestation of *Triatoma* (Bustamante-Zamora et al. 2015, Lima-Cordón et al. 2018), so it comes as no surprise

that the presence of biological factors such as humans, dogs, chickens, rodents, pigs, and livestock being kept in and around houses influences blood meal composition of these vectors. All possible domestic blood sources were present in the majority of houses, and LC-MS/MS identified these blood sources in the vector. In few instances LC-MS/MS detected a blood source not present in the household (Figure 4.5). For example, a dog or chicken blood meal was detected in a household that was not recorded to own either species (Table 4.1, Figure 4.5). However, due to the close proximity of many of these houses and the fact that dogs and some birds are free-ranging, it is not surprising a vector would contain blood meals from various sources nearby. Occasionally, LC-MS/MS detected a rodent while there was no visible presence of rodents in a home at the time of survey (Table 4.1, Figure 4.5). This could have simply been attributed to visually missing a trace of a rodent while the survey was conducted. In few instances, LC-MS/MS detected a rodent in a specimen, but we were not able to identify the blood meal to species level (i.e., Rodentia/Cricetidae or Sciuridae). This is most likely due to the rodent blood meal only being present in low abundance, or from the vector feeding on the rodent quite some time ago, possibly even during an earlier instar. In addition, there may also be a local species of rodent that does not have a hemoglobin sequence readily available in GenBank. Rodent presence has been known to be associated with persistent intra-domiciliary *T. dimidiata* infestation in Guatemala, and therefore rodent control could complement vector control programs (Bustamante et al. 2014, De Urioste-Stone et al. 2015). However, from a management

perspective, knowing vectors are feeding on rodents can lead to sound vector management decisions to eliminate rodents from houses.

Although from a few samples we were not able to detect the species level with stringent filtering or were not able to resolve to a particular rodent species, this does not deter from the ability to make vector management decisions based on the taxonomic order information of blood sources identified. Our multiple blood meal detection pipeline is able to classify multiple blood meals by order, which is often enough to determine what a vector is feeding on based on the habitat where the vector was collected and the ecology and behavior of the local insect species. This vector sample set from Guatemala had very low occurrence of sylvatic vector blood meals (2 % LC-MS/MS identified blood meals, and 7 % opossum blood meals in context of the overall study in Guatemala (Lima-Cordón et al. 2018)). Analyzing specimens from exclusively sylvatic habitats may lead to adjustments in our blood meal detection pipeline, and is something we have yet to explore. However, LC-MS/MS clearly identified the sylvatic opossum blood meal present in this data set.

Not only did hemoglobin-based LC-MS/MS detect a higher percentage of blood meals overall compared to DNA, but also within individual specimens (Figure 4.3, Figure 4.4). There was evidence for vectors having taken in multiple blood meals in 46 % of all specimens analyzed in comparison to only 8 % of vectors by DNA, and 0.32 % in context of the entire study (Lima-Cordón et al. 2018). Our blood source identification pipeline allowed us to confidentially identify multiple blood meals in this data set, which we confirmed by using control synthetic AQUA peptides. The presence

and absence of native peptides of target species confirmed the presence or absence of a particular blood meal source. For two samples, we did not initially identify a blood meal to species, yet with relaxing our high stringency peptide filter, we were able to identify both of these samples to the species level (Supplementary Table 4.6). While decreasing stringency can bring blood meal source numbers close to recently identified blood meals by genomics (Dumonteil et al. 2018), it can decrease the accuracy of peptide detection. Therefore, as our goal is to identify peptide sequences with utmost accuracy in order to detect different blood meal orders and even species with high accuracy, stringent peptide filtering is preferred for this task. Even with our stringent filtering and cut-offs, additional peptide and taxonomic information are not lost, and can be further explored. For example, peptides and orders falling into the 3 categories of E.2 in Figure 4.1 can easily be evaluated to validate additional minor blood meals. On a case by case basis, peptide filters can be made less stringent to identify lower quality blood meals, with the knowledge that species-specific identification needs to be examined carefully under lower stringency conditions. Even with our stringent filtering and major blood meal cut-offs, we detected 11 different blood meal sources in these vectors, a number comparable to similar studies (Stevens et al. 2012, Waleckx et al. 2014, Dumonteil et al. 2018, Lima-Cordón et al. 2018).

A great attribute of our protein-based method is the ability of LC-MS/MS to fill the gap of previously unidentifiable blood meals. We have previously shown that LC-MS/MS identified blood meals at least 4 weeks post-feeding and up to 12 weeks post-molting (Keller et al. 2018). Again here, LC-MS/MS identified blood meals in

specimens that did not have a blood meal detected with DNA-based methods (Keller et al. 2017, Keller et al. 2018). Another critical attribute of using mass spectrometry for blood meal analysis is the reasonable price tag. We have previously shown that samples can be run for as low as $4.75-$10, depending on self-run platforms versus sending samples to a core facility (Önder et al. 2013, Keller et al. 2017, Keller et al. 2018). Cost for DNA analysis by PCR include DNA extraction kits and PCR reagents and are comparable to preparing samples for LC-MS/MS. A major advantage of LC-MS/MS is the ability to determine multiple blood sources with a single LC-MS/MS run (Laskay et al. 2013), rather than single PCR reactions. In addition, cloning kits and cells (approximately $500 for 20 samples) are costly and DNA sequencing is comparable to outsourcing LC-MS/MS samples. Although genomics has the potential to identify multiple bloodmeal sources in a single specimen, costs can be over $50 per sample, require a large amount of high-quality DNA, and require additional reactions before genomic data is obtained (Kieran et al. 2017, Dumonteil et al. 2018). While top-of-the-line mass spectrometers may not be available in all Latin American countries, we show in this study that our methods and pipeline are easily tailored to identify blood meals using lower resolution mass spectrometers, such as the LTQ instrument tested here. Proteomic resources are available, especially in countries such as Brazil (Padrón and Domont 2014), and this opens the door to using LC-MS/MS-based blood meal identification in countries where Chagas disease is endemic, and resources are limited.

While presence/absence data of AQUA and native peptides worked well for identifying multiple blood meals and confirming our blood source pipeline, ideally

quantification of hemoglobin peptides can lead to identifying blood meal species not currently in our database. We can do this by not only quantifying general hemoglobin peptides, but also by quantifying species-specific peptides (Keller et al. 2017). This concept, however, will need further experimentation as some peptides decay at varying rates over time as the insect vector digests the blood meal (Laskay et al. 2013). Since species-specific peptides vary in location along the hemoglobin molecule, this can influence the amount of peptide identified. Therefore, we present here the first step in this endeavor by quantifying highly conserved regions of hemoglobin, a proxy for total amount of hemoglobin identified within an insect (Figure 4.6). Using spectral count as an indirect quantification method for a particular protein in a sample has been previously shown (Lundgren et al. 2010), and as expected, the number of hemoglobin peptides identified in a sample increased with the amount of hemoglobin. Knowing the amount of hemoglobin in a vector can help determine the quality of a blood meal and can aid in deciding if peptide filtering needs to be of lower stringency to increase the odds of identifying a blood meal in low quality specimens. This also opens the door for further exploration of species-specific AQUA peptide studies.

Mass spectrometry-based studies are on the rise and have a wide application, including for multi-faceted ecological studies like the one presented here. Although techniques like genomics may identify parasite burden in addition to blood meal composition, LC-MS/MS has the potential to be further applied for identifying *T. cruzi* parasite within a vector in future studies. Nevertheless, hemoglobin peptide-based LC-MS/MS allowed us to identify single and multiple blood meals in Triatomine insect

vectors with high confidence using a blood meal detection pipeline. Synthetic AQUA peptides allowed us to generally quantify the amount of hemoglobin in a vector, a first step in using AQUA peptides for further blood meal composition studies to include identifying blood meal species not present in our underlying database. Accurate blood meal detection at the order level, and even the species level identified here, can lead to sound vector management decisions and enhanced Ecohealth information for large-scale Ecohealth projects in Latin America.

4.6 Acknowledgements

We thank all collaborators at LENAP for insect sample collection and initial processing and Kristiaan Finstad and the UVM Stein Lab for providing a mouse blood control sample. Thank you to Alan Howard for valuable input on automating the blood meal pipeline. Thank you to Jeffrey Knott and Jason Reynolds at Cell Signaling Technology for synthesizing the stable isotope-containing peptides.

4.7 Figures and Tables

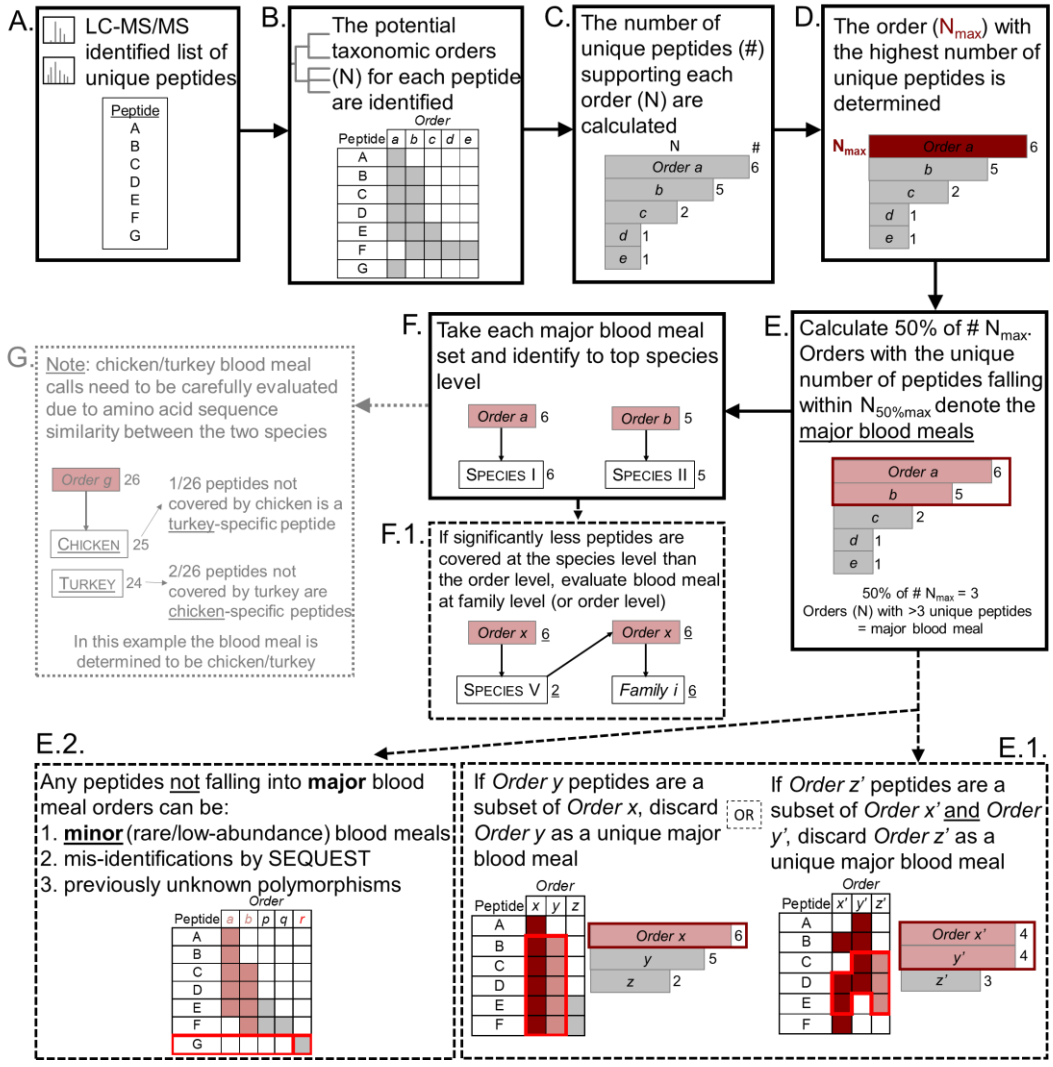

Figure 4.1: Workflow of updated pipeline for taxonomic evaluation of single and multiple blood meals identified by LC-MS/MS. (A) LC-MS/MS identifies hemoglobin peptides, from which a list of unique peptides is created. (B) Using taxonomy information from GenBank, the potential taxonomic orders are identified for each peptide, and (C) subsequently indexed- the number of unique peptides matching each order is recorded. The order with the highest number of unique peptides matching it (N_{max}) is identified (D) and (E) divided by 2 in order to determine the major blood meal cut-off. All orders with the number of unique peptides falling within 50% of N_{max} are designated as major blood meals. Orders are only considered as unique major blood meals if unique peptides are not subsets of another major blood meal (E.1). If any peptides do not fall into the major blood meal orders, they can be classified into the categories found in (E.2). Each major blood meal set is then identified to species (F), and to (F.1) family level if considerably less peptides are identified between order and species, or the order level if considerably less peptides are identified at the family level. (G) refers to a special case for chicken/turkey blood meals where hemoglobin amino acid sequences are highly similar between the two species.

Table 4.1: *Triatoma dimidiata* and *Triatoma nitida* sample information of specimens collected in Jutiapa, Guatemala in 2011 as part of a large-scale collaborative Ecohealth project. [a-e] designates specimens collected from the same individual house. All other vectors were collected from individual houses. * samples were not included in Lima-Cordón et al. 2018.

Sample ID							**Ecotope**	**Stage**	**Visible blood meal**	**Survey responses for household where specimen collected** .=not data available					
T. nitida and * samples not included in Lima-Cordón et al. 2018		Species of *Triatoma*	**Country**	**Department**	**Village** (Location)	**Subvillage**	Place where vector was collected	age of vector at collection	present/ absent/ unknown	# humans living in house	# dogs	# birds	# pigs	# beasts/ livestock	traces/ presence of rodents in house 1=yes 0=no
*Jkop_001	TPG_007	*dimidiata*	Guatemala	Chiquimula	Olopa	Amatillo	Intradomicile	V	unknown	.	.	.	.	.	.
*Jkop_002	TPG_011	*dimidiata*	Guatemala	Chiquimula	Olopa	Amatillo	Intradomicile	V	unknown	.	.	.	.	.	.
*Jkop_003	TPG_012	*dimidiata*	Guatemala	Chiquimula	Olopa	Amatillo	Intradomicile	V	unknown	.	.	.	.	.	.
Jkop_004	TPG0028[e]	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	female	unknown	8	4	0	0	0	1
Jkop_005	TPG0030[e]	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	III	unknown	8	4	0	0	0	1
Jkop_006	TPG0038	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	male	unknown	.	.	.	.	.	.
Jkop_007	TPG0044	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	V	present	3	0	2	0	0	0
*Jkop_008	TPG_067	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	unknown	unknown	.	.	.	.	.	.
Jkop_009	TPG0075	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	2	12	5	0	0	1
Jkop_011	TPG_086	*nitida*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	III	unknown	4	2	12	0	0	1
Jkop_012	TPG_087	*nitida*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	male	unknown	6	4	5	0	1	1
Jkop_013	TPG_088	*nitida*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	male	unknown	8	1	4	0	0	1
Jkop_015	TPG_092	*nitida*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	female	unknown	6	0	6	0	0	1
Jkop_016	TPG_097[b]	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	11	3	4	0	0	1
Jkop_017	TPG_100[a]	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	4	6	2	0	1	1
Jkop_018	TPG_101[a]	*nitida*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	4	6	2	0	1	1
Jkop_019	TPG_104[a]	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	4	6	2	0	1	1
Jkop_020	TPG_178	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	5	1	15	0	0	0
Jkop_021	TPG_158[d]	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	V	present	5	1	18	0	0	1
Jkop_022	TPG_168	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Peridomicile	male	present	5	0	8	0	0	1
Jkop_023	TPG_186	*nitida*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	male	unknown	3	5	35	0	0	1
Jkop_024	TPG_238[a]	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	III	unknown	4	6	2	0	1	1
Jkop_025	TPG_336	*dimidiata*	Guatemala	Chiquimula	Olopa	El Cerrón	Intradomicile	V	unknown	6	2	8	0	0	1
Jkop_027	TPG_630	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	female	present	7	2	8	0	0	1
Jkop_028	TPG_670	*nitida*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	male	unknown	9	2	12	0	0	1
*Jkop_029	TPG_016	*dimidiata*	Guatemala	Chiquimula	Olopa	Amatillo	Intradomicile	III	present	10	0	11	1	0	1
Jkop_030	TPG0045	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	male	absent	3	3	1	0	0	0
Jkop_031	TPG0102	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	V	present	4	6	2	0	1	1
Jkop_032	TPG0159[d]	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	V	present	5	1	18	0	0	1
Jkop_033	TPG0244[b]	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	III	absent	11	3	4	0	0	1
Jkop_034	TPG0263	*dimidiata*	Guatemala	Chiquimula	Olopa	El Guayabo	Intradomicile	III	present	.	.	.	.	.	.
Jkop_035	TPG0288	*dimidiata*	Guatemala	Chiquimula	Olopa	El Cerrón	Intradomicile	III	absent	10	1	1	0	0	1
Jkop_036	TPG0302	*dimidiata*	Guatemala	Chiquimula	Olopa	El Cerrón	Intradomicile	III	present	2	0	0	0	0	1
Jkop_037	TPG0334	*dimidiata*	Guatemala	Chiquimula	Olopa	El Cerrón	Intradomicile	III	present	6	2	8	0	0	1
Jkop_038	TPG0381	*dimidiata*	Guatemala	Chiquimula	Olopa	Paternito	Peridomicile	female	absent	7	2	8	0	0	1
Jkop_039	TPG0385	*dimidiata*	Guatemala	Chiquimula	Olopa	Paternito	Intradomicile	III	absent	4	3	10	0	1	1
Jkop_040	TPG0412[c]	*dimidiata*	Guatemala	Chiquimula	Olopa	Paternito	Intradomicile	III	present	7	2	8	0	0	1
Jkop_041	TPG0415[c]	*dimidiata*	Guatemala	Chiquimula	Olopa	Paternito	Intradomicile	IV	present	7	2	8	0	0	1

Jkop_042	TPG0471	*dimidiata*	Guatemala	Chiquimula	Olopa	Paternito	Intradomicile	IV	present	9	2	10	0	0	1
Jkop_043	TPG0585	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	unknown	V	present	4	0	6	0	0	0
Jkop_044	TPG0596	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	IV	present	4	2	22	0	0	1
Jkop_045	TPG0603	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	IV	present	6	1	21	0	0	1
Jkop_046	TPG0635	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	V	present	.	.	.	.	.	.
Jkop_047	TPG0679	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Peridomicile	male	absent	6	1	0	0	0	1
Jkop_048	TPG0703	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	V	present	5	2	0	0	0	1
Jkop_049	TPG0749	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	IV	present	8	2	0	0	0	1
Jkop_050	TPG0771	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	male	present	6	0	0	0	0	1
Jkop_051	TPG0807	*dimidiata*	Guatemala	Chiquimula	Olopa	El Cerrón	Peridomicile	III	present	4	2	0	2	0	1
Jkop_052	TPG0814	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	male	absent	6	2	0	0	0	1
Jkop_053	TPG0717	*dimidiata*	Guatemala	Chiquimula	Olopa	La Prensa	Intradomicile	IV	present	4	1	0	0	0	1

Table 4.2: Synthetic AQUA peptides designed for determination of multiple blood meals and quantification. Synthetically modified heavier amino acids are marked as bold and underlined. Two known polymorphisms of the hemoglobin beta position 31 peptide exist, LLIVYPWTQR and LLVVYPWTQR. One of these two variants occur in over 639 species (17 Jan 2018).

Name	ID	Sequence	Species	Molecular Weight	Dominant charge state (m/z)	(2+) Light sequence (m/z)
M14-28011	"dog"	FFAAVSTV**l**TSK	*Canis lupus familiaris* (dog), 5 lemur spp.	1276.715	638.865	635.856
M14-28012	"chicken/turkey"	FLSAVSAV**l**AEK	*Gallus gallus* (chicken), *Meleagris gallopavo* (turkey), 4 other galliformes	1240.715	620.865	617.856
M14-28013	"birds"	VLTSFGDA**v**K	15 Avian orders, including duck, songbird, chicken, turkey, 7 reptiles, 31 mammals	1041.574	521.295	518.787
M14-28016	"human"	FFESFGDLSTPDAVMGN**p**K	*Homo sapiens* (human), 7 other primates	2063.955	1032.485	1029.978
M14-3888	general	LLIVY**p**WTQR	General peptides of highly conserved regions of hemoglobin matching over 639 species	1293.748	647.882	644.875
M14-3889	general	LLVVY**p**WTQR		1279.733	640.874	637.867

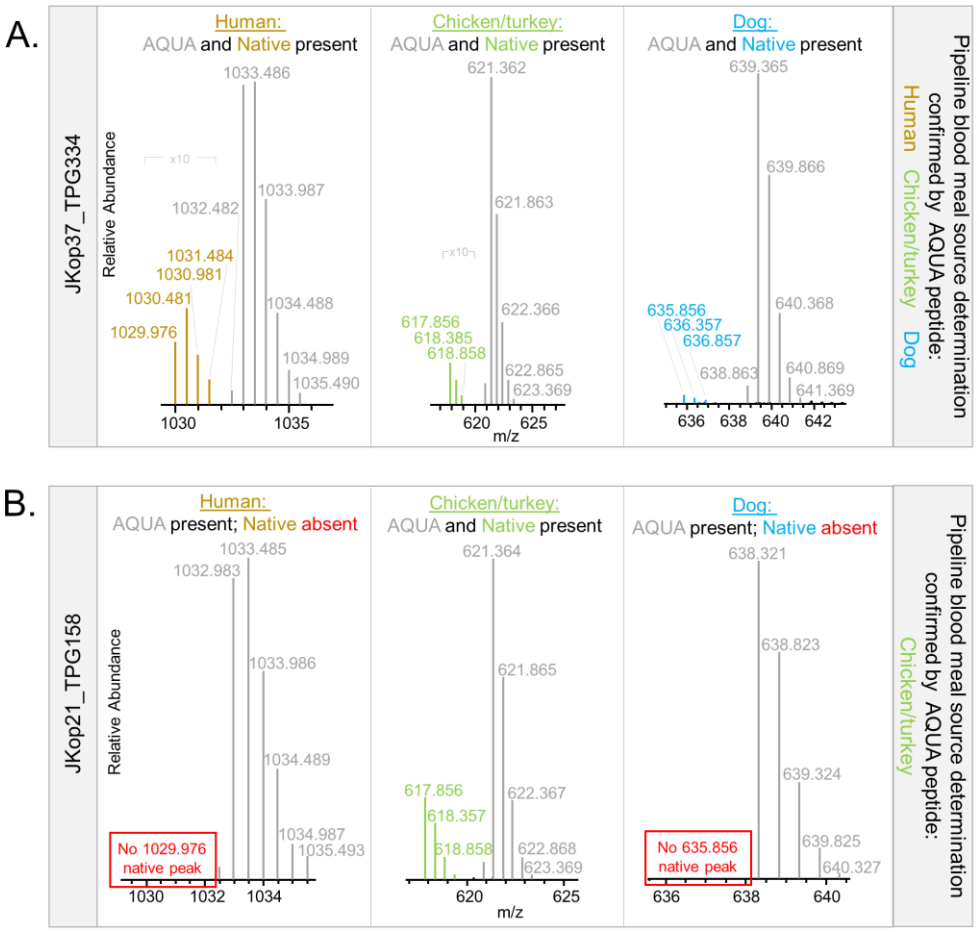

Figure 4.2: AQUA and native peptide presence and absence validates blood meal pipeline. Two representative samples are shown, a multiple and single blood meal, where (A) shows the presence of dog, human, and chicken/turkey ms1 isotopic native peaks, which agrees with the blood meals identified by the pipeline being dog, human, and chicken/turkey. (B) shows a single blood meal of chicken/turkey, also identified by the pipeline, with lacking dog and human native peptide peaks.

Table 4.3: Presence and absence data for ms1 monoisotopic peaks corresponds to blood meal pipeline identified blood meals. + present; • absent; * a native peptide was not present- see text for further details.

AQUA peptide		Sample ID										
Name	Species	JKop53_ TPG717	JKop51_ TPG807	JKop37_ TPG334	JKop34_ TPG263	JKop30_ TPG045	JKop27_ TPG630	JKop23_ TPG186*	JKop22_ TPG168	JKop21_ TPG158	JKop19_ TPG104*	JKop09_ TPG075*
M14-28011	"dog"	.	.	+	.	.	.	.	.	.	+	+
M14-28012	"chicken/turkey"	+	.	+	.	.	+	.	+	+	.	.
M14-28013	"birds"	+	.	+	.	+	+	+	+	+	.	.
M14-28016	"human"	.	.	+	+	.	.	.	.	.	.	.
LC-MS/MS pipeline identified blood meal		**chicken/ turkey**	**pig, rodentia**	**dog, chicken/ turkey, human**	**human, opossum**	**duck**	**chicken/ turkey**	**turkey/ chicken**	**chicken/ turkey**	**chicken/ turkey**	**dog, rodentia, human**	**dog, human, rodentia**

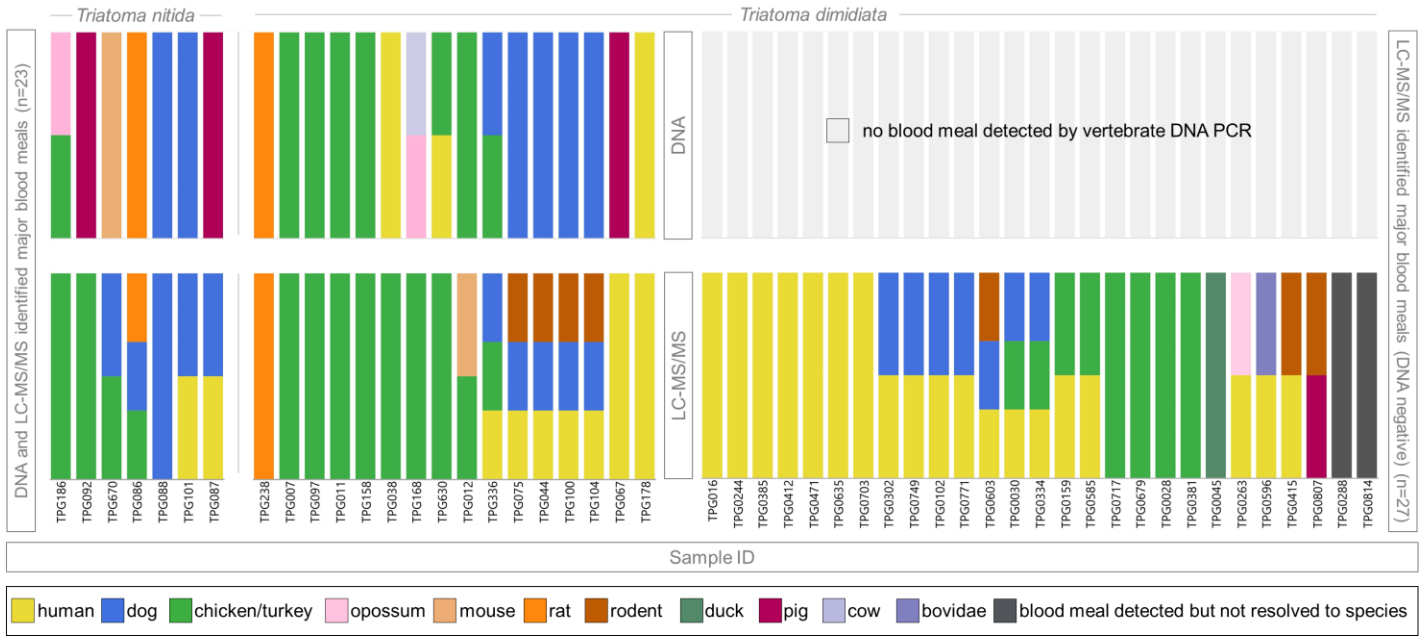

Figure 4.3: DNA and LC-MS/MS identified blood meal sources in 50 Triatomine vectors from Jutiapa, Guatemala. DNA-identified blood meals were only identified to class level for birds and are indicated by green color in comparison to LC-MS/MS that identified to species level (chicken/turkey). JKop06_TPG038 was identified as chicken/turkey/francolin (shown in figure as chicken/turkey). JKop28-TPG670 was identified to 6 Canidae (shown in figure as dog).

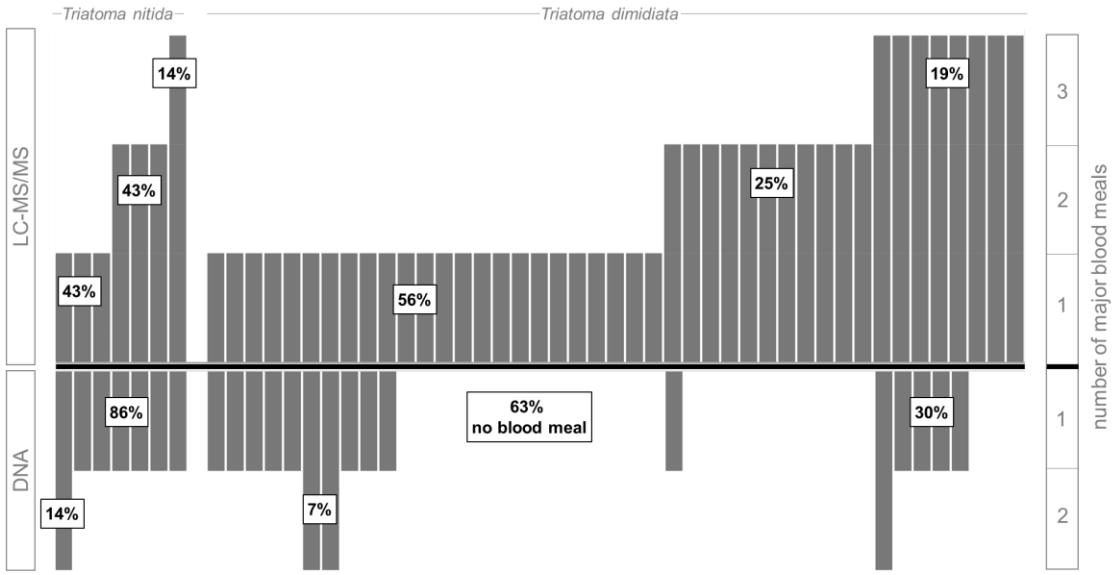

Figure 4.4: Number of major blood meal sources detected within an individual specimen with DNA and LC-MS/MS in 50 Triatomines from Guatemala. Percentages indicate how many specimens had zero, one, two, or three blood meals by method (DNA or LC-MS/MS) and by vector species (*T. nitida*, n=7; *T. dimidiata* n=43).

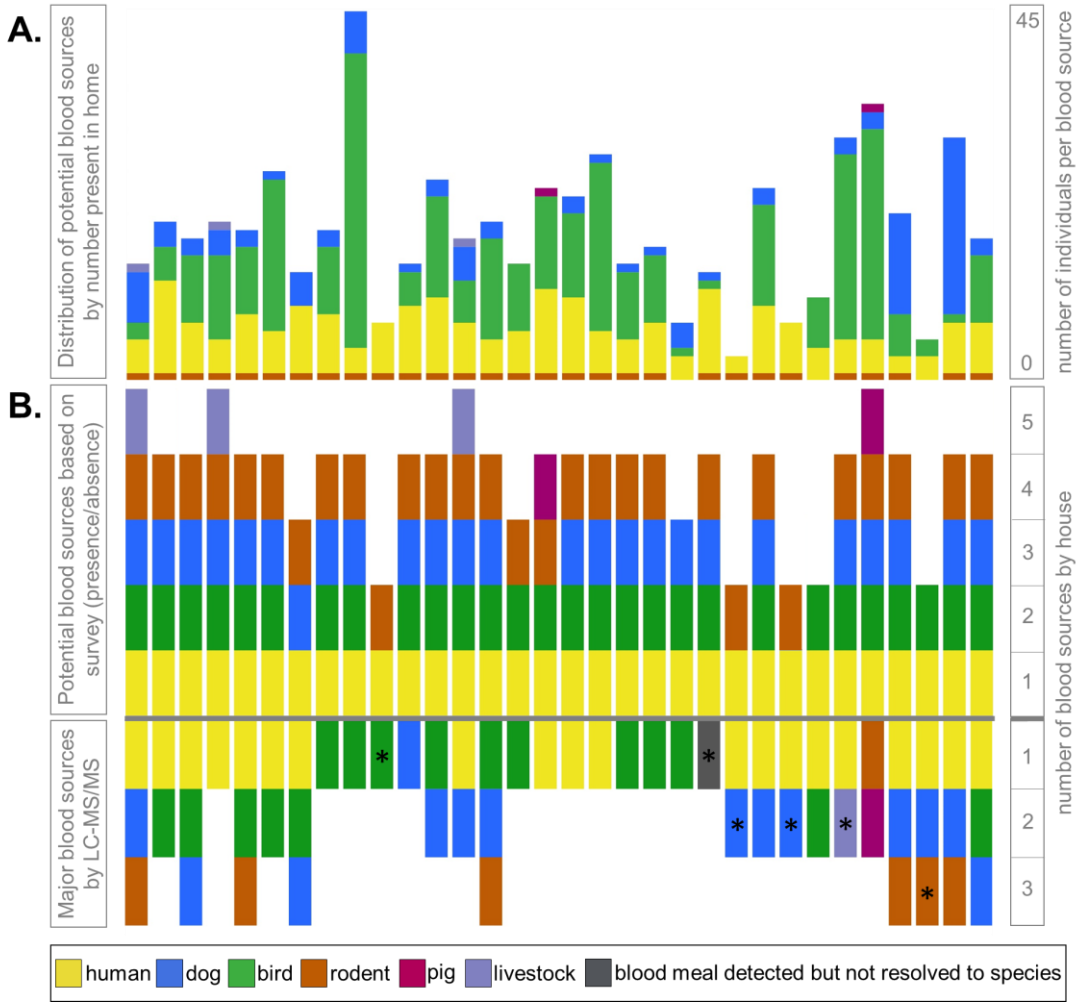

Figure 4.5: Potential blood sources per household based on pre-intervention survey data. (A) The distribution of blood meal sources per household is shown based on survey data, including the number of inhabitants per house (human blood source). Presence or trace of rodents is shown as presence/absence. (B) Potential household blood sources are compared to LC-MS/MS identified major blood meals. (*) LC-MS/MS detected a blood source not present in household. Only houses with complete survey data are shown.

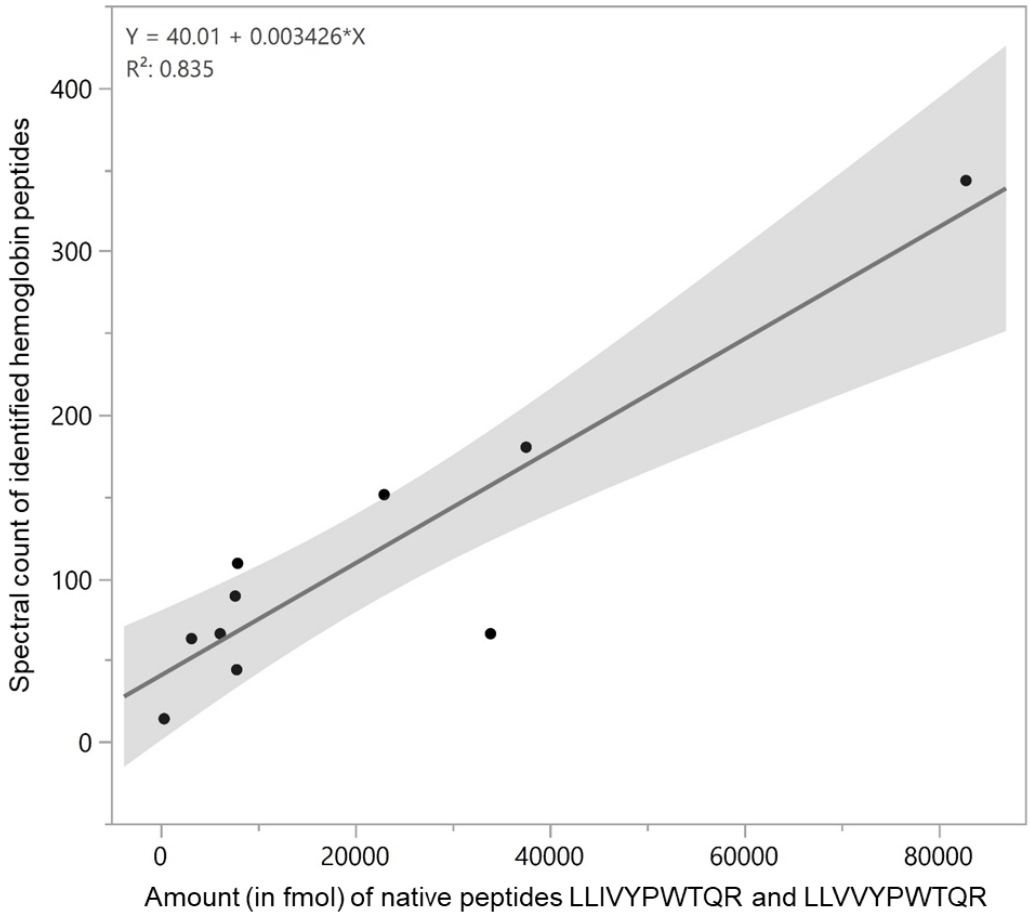

Figure 4.6: The amount of hemoglobin in a vector increased with spectral count. Spectral count is the number of hemoglobin peptides identified in a sample. By fitting a linear regression line (R2 = 0.835), fmol amount of native peptide increased with the number of hemoglobin peptides identified.

Supplementary Table 4.1: Details of PCR conditions used in Lima-Córdon et al. 2018 showing primer sequences, element targeted, and reference.

Supplementary Table 4.2: Taxonomic adjustments of entries with missing taxa information in GenBank. Classes and Orders in italics were blank in GenBank and populated manually for the purpose of complete taxonomic assignment of peptides.

Supplementary Table 4.3: Blood meal source detection pipeline information showing blood meal species assignment with supporting data.

Supplementary Table 4.4: Blood meal source detection comparison between DNA and LC-MS/MS identified blood meal sources.

Supplementary Table 4.5: Blood meal source detection of a single and multiple Triatominae specimen blood meal and mouse blood control comparing two LC-MS/MS-based mass spectrometry instruments: Orbitrap and LTQ.

Supplementary Table 4.6: Blood meal source detection of a subset of Triatominae specimen evaluating high and lower stringency peptide filters of SEQUES-identified peptides. For more detail of high stringency peptide support see Supplementary Table 4.3.

4.8 References

Ballif, B. A., G. R. Carey, S. R. Sunyaev, and S. P. Gygi. 2008. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. Journal of Proteome Research 7:311-318.

Bustamante-Zamora, D. M. B., M. M. Hernández, N. Torres, C. Zúniga, W. Sosa, V. de Abrego, and M. C. M. Escobar. 2015. Information to act: household characteristics are predictors of domestic infestation with the Chagas vector Triatoma dimidiata in Central America. The American Journal of Tropical Medicine and Hygiene 93:97-107.

Bustamante, D. M., S. M. De Urioste-Stone, J. G. Juárez, and P. M. Pennington. 2014. Ecological, social and biological risk factors for continued Trypanosoma cruzi transmission by Triatoma dimidiata in Guatemala. PLoS One 9:e104599.

Cecere, M. C., G. M. Vazquez-Prokopec, R. E. Gürtler, and U. Kitron. 2006. Reinfestation sources for Chagas disease vector, Triatoma infestans, Argentina. Emerging Infectious Diseases 12:1096.

Collini, M., H. Hauffe, S. Masseglia, F. Albonico, D. Arnoldi, X. Bailly, E. Bard, M. Galan, C. Rossi, and V. Tagliapietra. 2015. Identifying the last bloodmeal of questing wood tick nymphs (Ixodes ricinus L.) by DNA amplification: three approaches tested. Page Poster 2.18 in GERI 2015: Genes, Ecosystems and Risk of Infection.

De Urioste-Stone, S. M., P. M. Pennington, E. Pellecer, T. M. Aguilar, G. Samayoa, H. D. Perdomo, H. Enriquez, and J. G. Juárez. 2015. Development of a community-based intervention for the control of Chagas disease based on peridomestic animal management: an eco-bio-social perspective. Transactions of the Royal Society of Tropical Medicine and Hygiene 109:159-167.

Dumonteil, E., M.-J. Ramirez-Sierra, S. Pérez-Carrillo, C. Teh-Poot, C. Herrera, S. Gourbière, and E. Waleckx. 2018. Detailed ecological associations of triatomines revealed by metabarcoding and next-generation sequencing: implications for triatomine behavior and Trypanosoma cruzi transmission cycles. Scientific reports 8:4140.

Georgieva, A. Y., E. R. L. Gordon, and C. Weirauch. 2017. Sylvatic host associations of Triatominae and implications for Chagas disease reservoirs: a review and new host records based on archival specimens. PeerJ 5:e3826.

Hashimoto, K., C. Cordon-Rosales, R. Trampe, and M. Kawabata. 2006. Impact of single and multiple residual sprayings of pyrethroid insecticides against Triatoma dimidiata (Reduviiade; Triatominae), the principal vector of Chagas disease in Jutiapa, Guatemala. American Journal of Tropical Medicine and Hygiene 75:226.

Justi, S. A., and C. Galvão. 2017. The Evolutionary Origin of Diversity in Chagas Disease Vectors. Trends in parasitology 33:42-52.

Keller, J. I., B. A. Ballif, R. M. S. Clair, J. J. Vincent, M. C. Monroy, and L. Stevens. 2017. Chagas disease vector blood meal sources identified by protein mass spectrometry. PLoS One 12:e0189647.

Keller, J. I., J. O. Schmidt, A. M. Schmoker, B. A. Ballif, and L. Stevens. 2018. Protein mass spectrometry extends temporal blood meal detection over polymerase chain reaction in mouse-fed Chagas disease vectors. Memórias do Instituto Oswaldo Cruz 113.

Kent, R. J. 2009. Molecular methods for arthropod bloodmeal identification and applications to ecological and vector-borne disease studies. Molecular ecology resources 9:4-18.

Kieran, T. J., N. L. Gottdenker, C. P. Varian, A. Saldaña, N. Means, D. Owens, J. E. Calzada, and T. C. Glenn. 2017. Blood meal source characterization using Illumina Sequencing in the Chagas Disease vector Rhodnius pallescens (Hemiptera: Reduviidae) in Panamá. Journal of medical entomology 54:1786-1789.

Klotz, S. A., J. O. Schmidt, P. L. Dorn, C. Ivanyi, K. R. Sullivan, and L. Stevens. 2014. Free-roaming kissing bugs, vectors of Chagas disease, feed often on humans in the Southwest. The American Journal of Medicine 127:421-426.

Laskay, Ü. A., L. Breci, I.-M. E. Vilcins, G. Dietrich, A. G. Barbour, J. Piesman, and V. H. Wysocki. 2013. Survival of host blood proteins in Ixodes scapularis (Acari: Ixodidae) ticks: a time course study. Journal of medical entomology 50:1282-1290.

Lima-Cordón, R. A., L. Stevens, E. S. Ortíz, G. A. Rodas, V. Abrego, C. Zúñiga, and C. Monroy. 2018. Implementation science: Epidemiology and feeding profiles of the Chagas vector Triatoma dimidiata prior to Ecohealth intervention for three locations in Central America. PLOS Negl Trop Dis accepted

Logue, K., J. B. Keven, M. V. Cannon, L. Reimer, P. Siba, E. D. Walker, P. A. Zimmerman, and D. Serre. 2016. Unbiased characterization of Anopheles mosquito blood meals by targeted high-throughput sequencing. PLoS neglected tropical diseases 10:e0004512.

Lucero, D. E., L. A. Morrissey, D. M. Rizzo, A. Rodas, R. Garnica, L. Stevens, D. M. Bustamante, and M. C. Monroy. 2013. Ecohealth interventions limit Triatomine reinfestation following insecticide spraying in La Brea, Guatemala. The American Journal of Tropical Medicine and Hygiene 88:630-637.

Lucero, D. E., W. Ribera, J. C. Pizarro, C. Plaza, L. W. Gordon, R. Pena, Jr., L. A. Morrissey, D. M. Rizzo, and L. Stevens. 2014. Sources of Blood Meals of Sylvatic Triatoma guasayana near Zurima, Bolivia, Assayed with qPCR and 12S Cloning. PLOS Negl Trop Dis 8:e3365.

Lundgren, D. H., S.-I. Hwang, L. Wu, and D. K. Han. 2010. Role of spectral counting in quantitative proteomics. Expert review of proteomics 7:39-53.

Nakagawa, J., C. Cordón-Rosales, J. Juárez, C. Itzep, and T. Nonami. 2003. Impact of residual spraying on Rhodnius prolixus and Triatoma dimidiata in the department of Zacapa in Guatemala. Memórias do Instituto Oswaldo Cruz 98:277-282.

Niare, S., J.-M. Berenger, C. Dieme, O. Doumbo, D. Raoult, P. Parola, and L. Almeras. 2016. Identification of blood meal sources in the main African malaria mosquito vector by MALDI-TOF MS. Malaria journal 15:1.

Önder, Ö., W. Shao, B. D. Kemps, H. Lam, and D. Brisson. 2013. Identifying sources of tick blood meals using unidentified tandem mass spectral libraries. Nature communications 4:1746.

Orantes, L. C., C. Monroy, P. L. Dorn, L. Stevens, D. M. Rizzo, L. Morrissey, J. P. Hanley, A. G. Rodas, B. Richards, and K. F. Wallin. 2018. Uncovering vector, parasite, blood meal and microbiome patterns from mixed-DNA specimens of the Chagas disease vector Triatoma dimidiata. PLoS neglected tropical diseases 12:e0006730.

Padrón, G., and G. B. Domont. 2014. Two decades of proteomics in Latin America: a personal view. Journal of proteomics 107:83-92.

Pellecer, M. J., P. L. Dorn, D. M. Bustamante, A. Rodas, and M. C. Monroy. 2013. Vector blood meals are an early indicator of the effectiveness of the Ecohealth approach in halting Chagas transmission in Guatemala. The American Journal of Tropical Medicine and Hygiene 88:638-644.

Pizarro, J. C., D. Lucero, and L. Stevens. 2007. A method for the identification of guinea pig blood meal in the Chagas disease vector, Triatoma infestans. Kinetoplastid biology and disease 6:1.

Pizarro, J. C., and L. Stevens. 2008. A new method for forensic DNA analysis of the blood meal in Chagas disease vectors demonstrated using Triatoma infestans from Chuquisaca, Bolivia. PLoS One 3:e3585.

Rabinovich, J. E., U. D. Kitron, Y. Obed, M. Yoshioka, N. Gottdenker, and L. F. Chaves. 2011. Ecological patterns of blood-feeding by kissing-bugs (Hemiptera: Reduviidae: Triatominae). Memórias do Instituto Oswaldo Cruz 106:479-494.

Stevens, L., P. L. Dorn, J. Hobson, N. M. de la Rua, D. E. Lucero, J. H. Klotz, J. O. Schmidt, and S. A. Klotz. 2012. Vector blood meals and Chagas disease transmission potential, United States. Emerging Infectious Diseases 18:646-649.

Waleckx, E., J. Camara-Mejia, M. J. Ramirez-Sierra, V. Cruz-Chan, M. Rosado-Vallado, S. Vazquez-Narvaez, R. Najera-Vazquez, S. Gourbière, and E. Dumonteil. 2015. An innovative ecohealth intervention for Chagas disease vector control in Yucatan, Mexico. Transactions of the Royal Society of Tropical Medicine and Hygiene 109:143-149.

Waleckx, E., J. Suarez, B. Richards, and P. L. Dorn. 2014. Triatoma sanguisuga blood meals and potential for Chagas disease, Louisiana, USA. Emerg Infect Dis 20:2141-2143.

World Health Organization. 2016. Chagas disease (American trypanosomiasis). Fact sheet N 340 (2015). World Health Organization.

World Health Organization, and UNICEF. 2017. Global vector control response 2017-2030.

CHAPTER 5: FUTURE DIRECTION AND BROADER IMPLICATIONS

Proteomics studies are on the rise and mass spectrometry has become an important analytical tool in many disciplines to aid in the identification of proteins. Constantly evolving and improving, even with minimal processing of samples, many instruments are capable of identifying thousands of protein peptide sequences in under an hour. The interdisciplinary nature of a systematic tool like protein mass spectrometry allows for the application to intricate systems, such as presented in this body of work. Protein mass spectrometry allowed us to apply an innovate approach to the detection of blood meal sources in the complex vector-borne Chagas disease system. Although few other studies have evaluated protein mass spectrometry-based techniques in vector-borne disease systems, our studies present the first comprehensive tool that can be applied to field-collected specimens, as well as experimental populations. While our harvested hemoglobin peptide data requires expert processing through a blood meal detection pipeline, we are able to determine blood meals in a straight-forward manner that can easily be applied to other disease systems.

The World Health organization classifies the Triatomine vectors of Chagas disease as vectors of significance and has identified vector management as a priority research area (World Health Organization 2012). The data and methodology presented in these studies aid in this endeavor and directly impact the field of reducing the Chagas disease burden. We present the first studies using proteomic methods for blood meal identification in Chagas disease vectors and provide landmark advances in enhancing blood meal detection techniques as a whole. A second far reaching benefit

includes the direct application of the results presented in this body of work to Chagas disease prevention and vector management. The discoveries made in these studies are crucial in furthering our understanding of Chagas disease ecology and open the door to using proteomic techniques in Chagas endemic regions, suggesting further management implications.

We know that battling vector-borne diseases such as Chagas require innovative disease management techniques and that vector control is the most useful method to prevent Chagas disease in Latin America. Ecohealth (Lucero et al. 2013), an integrative, ecosystem-based approach using implementation science (Henao-Martínez et al. 2017) that factors the results of field-collected data into management decisions for affected communities, can aid in disease prevention. Not only does knowing vector blood meal prevalence provide valuable data for Ecohealth management, it can also help evaluate intervention measures to help make practical recommendations as shown in Pellecer et al. 2013. Since commonly used DNA-based blood meal detection is generally restricted to recent blood meals in field-collected vectors, LC-MS/MS can help fill this gap in knowledge and provide a more complete picture of what vectors are feeding on in the wild. In addition, there is the potential to identify not only the blood meal, but also the *T. cruzi* parasite in the insect vector using LC-MS/MS. This endeavor will need controlled experiments and additional bioinformatics resources. However, identifying infection status of the vector along with the vector blood meal is paramount in deciphering disease ecology and a priority for further research and method development. Coupling vector host choice with Chagas parasite infections status will

allow us to determine the effects of local vector populations on human health and the risks they pose.

Protein mass spectrometry is a versatile tool that can be applied to many facets of Ecohealth disease management and surveillance, not just in Chagas vectors. Outstanding characteristics of our LC-MS/MS-based approach are 1) the ability to detect multiple, complex blood meals within a single insect vector and in instances where DNA does not offer blood meal identification; 2) the ability to detect blood meals over long temporal scales, especially compared to DNA; and 3) the ability to fine-tune peptide filtering and detection pipeline cut-offs to account for differences in lower resolution instruments that may be available in limited-resource regions of the world, such as in Chagas endemic countries.

Although as blood meals decay our method has limited taxonomic resolution for older blood meals, it is still more powerful than DNA for determining blood meal source prevalence for disease management decisions and for influencing policy. For example, when evaluating blood meal sources for policy decisions, blood meals generally only need to be identified to the taxonomic order level. Knowing that vectors are feeding on gallinaceous birds (Galliformes, chickens and turkeys), for instance, is enough information to establish a policy that controls birds such as chickens and turkeys to curb contact between vectors and potential human hosts, even without detecting blood meals to the species level. On the other hand, when considering less well characterized environments, such as sylvatic habitats, knowing vector blood meals to the species level may be important to elucidate novel host species. Therefore,

variability in detection ability could influence deciphering some aspects of disease ecology.

However, when compared to a recent study examining novel blood meals in sylvatic vectors, our approach could potentially have identified these blood meals to species as hemoglobin sequences were readily available in GenBank for almost all of the new host associations (Georgieva et al. 2017). Although there is the potential for missing hemoglobin sequence data in GenBank of sylvatic, understudied animal species, we can often identify what vectors are feeding on to a higher taxonomic designation in these cases, i.e. order or family, though this requires expertise in data analysis. Nonetheless, these blood meal identifications can still lead to sound management decisions and understanding disease ecology based on knowing the ecology of the habitat where work is being conducted and local animal host populations. Factoring in the ecology and behavior of the vector and what is known of the region is part of the interdisciplinary nature of this research. Although our approach may be slightly conservative at times, it is highly accurate.

The interdisciplinary nature of this work allowed us to combine many different tools, from proteomics and genomics, to vector ecology and behavior. This can lead to the larger scale application of using LC-MS/MS-based techniques not only in the field for Chagas disease vectors, but also with other disease vectors, allowing us to address new questions in vector ecology.

Although Chagas disease is one of the most significant vector borne diseases and disproportionally affects people living in poor regions of the world, other

hematophagous vectors that cause diseases in more developed countries include mosquitos and ticks. Including in the United States, mosquito-borne diseases such as West Nile and Zika have dramatically risen in recent years, while Lyme disease caused by ticks is at its highest disease prevalence ever, and likely is underestimated (World Health Organization 2014). Although the aim of our studies was not to evaluate LC-MS/MS-based blood meal identification in all hematophagous vectors, the sensitivity of our LC-MS/MS-based approach makes it reasonable for the application to other disease vectors that are significantly smaller in overall size as compared to the Triatomines. For smaller vectors especially, which are routinely pooled when processed for blood meals, additional fine-tuning of data analysis methods may be required, but could potentially decrease overall processing time of samples.

In addition, similar LC-MS/MS detection techniques have already identified blood meals almost a year post-molting in ticks, a much smaller vector than the Triatomine vectors of Chagas disease (Laskay et al. 2013). Our approach would allow for application to field-collected ticks, something previous studies have not addressed. While ticks digest blood meals more slowly, in mosquitos the digestion of blood can take a few hours to a few days and is promoted by rapid degradation of host DNA, limiting blood meal detection efficiency (Martínez-de la Puente et al. 2013).

Different mass spectrometry-based approaches for blood meal detection have been employed in mosquitos (e.g., MALDI-TOF), with varying success (Niare et al. 2016). Various species of mosquitos can serve as vectors of malaria, leishmaniasis, dengue, yellow fever and other diseases (Lehane 2005). Therefore, knowing the host

choice of the mosquito vector can help identify different disease reservoirs and also determine vector competence (Logue et al. 2016). Just as with Chagas disease, knowing vector blood meals can help develop effective vector-control programs. Applying an LC-MS/MS-based blood meal identification method to mosquitos could therefore help elucidate previously unknown disease reservoirs and offers a new tool for evaluating mosquito blood meals. Applying the methodology presented in this body of work to a broader field is an important next step and an example of addressing similar questions in other disease systems.

With our detailed feeding profiles of various vectors, we are able to observe changes in the host species composition in congruence with changes in the environment. Global change, including climate change, can increase the vulnerability of people to vector-borne diseases (Sutherst 2004, Campbell-Lendrum et al. 2015). Insect vectors tend to have some flexibility in environmental habitat preference and host species, so changes in climate and environment can increase chances for human contact and disease transmission. In a time where environments are drastically changing, which often influences disease systems negatively, identifying vector blood meal hosts prevalence is vital for reducing human incidences of vector-borne diseases.

Vector-borne disease, and especially the Chagas disease burden in Latin America is a major one. Innovative research is critical to disease management and the direct application of our results to Chagas disease prevention, management, and surveillance has the power to dramatically change the lives of people living in endemic areas. In the future, studies such as the ones presented here will continue to be

increasingly more important, aiding in our understanding of vector-borne diseases and how to control them.

COMPREHENSIVE BIBLIOGRAPHY

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of molecular biology 215:403-410.

Aufderheide, A. C., W. Salo, M. Madden, J. Streitz, J. Buikstra, F. Guhl, B. Arriaza, C. Renier, L. E. Wittmers, Jr., G. Fornaciari, and M. Allison. 2004. A 9,000-year record of Chagas' disease. Proceedings of the National Academy of Sciences of the United States of America 101:2034-2039.

Ballif, B. A., G. R. Carey, S. R. Sunyaev, and S. P. Gygi. 2008. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. Journal of Proteome Research 7:311-318.

Belles, X. 2011. Origin and evolution of insect metamorphosis. Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd, Chichester, UK.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2009. GenBank. Nucleic Acids Res 37:D26-31.

Bern, C., S. Kjos, M. J. Yabsley, and S. P. Montgomery. 2011. Trypanosoma cruzi and Chagas' Disease in the United States. Clinical microbiology reviews 24:655-681.

Boischio, A., A. Sánchez, Z. Orosz, and D. Charron. 2009. Health and sustainable development: challenges and opportunities of ecosystem approaches in the prevention and control of dengue and Chagas disease. Cadernos de Saúde Pública 25:S149-S154.

Bosseno, M. F., L. S. Garcia, F. Baunaure, E. M. Gastelum, M. S. Gutierrez, F. L. Kasten, E. Dumonteil, and S. F. Breniere. 2006. Identification in triatomine vectors of feeding sources and Trypanosoma cruzi variants by heteroduplex assay and a multiplex miniexon polymerase chain reaction. The American Journal of Tropical Medicine and Hygiene 74:303-305.

Breniere, S. F., S. Pietrokovsky, E. M. Gastelum, M. F. Bosseno, M. M. Soto, A. Ouaissi, F. L. Kasten, and C. Wisnivesky-Colli. 2004. Feeding patterns of Triatoma longipennis Usinger (Hemiptera, Reduviidae) in peridomestic habitats of a rural community in Jalisco State, Mexico. Journal of medical entomology 41:1015-1020.

Buitrago, N. L. R., M. F. Bosseno, E. Waleckx, P. Brémond, P. Vidaurre, F. Zoveda, and S. F. Brenière. 2013. Risk of transmission of Trypanosoma cruzi by wild Triatoma infestans (Hemiptera: Reduviidae) in Bolivia supported by the detection of human blood meals. Infection, Genetics and Evolution 19:141-144.

Buitrago, R., M.-F. Bosseno, S. Depickère, E. Waleckx, R. Salas, C. Aliaga, C. Barnabé, and S. F. Brenière. 2016. Blood meal sources of wild and domestic Triatoma infestans (Hemiptera: Reduviidae) in Bolivia: connectivity between cycles of transmission of Trypanosoma cruzi. Parasites & vectors 9:1.

Buitrago, R., S. Depickère, M.-F. Bosseno, E. S. Patzi, E. Waleckx, R. Salas, C. Aliaga, and S. F. Brenière. 2012. Combination of cytochrome b heteroduplex-assay and sequencing for identification of triatomine blood meals. Infection, Genetics and Evolution 12:21-27.

Bustamante-Zamora, D. M. B., M. M. Hernández, N. Torres, C. Zúniga, W. Sosa, V. de Abrego, and M. C. M. Escobar. 2015. Information to act: household characteristics are predictors of domestic infestation with the Chagas vector Triatoma dimidiata in Central America. The American Journal of Tropical Medicine and Hygiene 93:97-107.

Bustamante, D. M., S. M. De Urioste-Stone, J. G. Juárez, and P. M. Pennington. 2014. Ecological, social and biological risk factors for continued Trypanosoma cruzi transmission by Triatoma dimidiata in Guatemala. PLoS One 9:e104599.

Calzada, J. E., V. Pineda, J. D. Garisto, F. Samudio, A. M. Santamaria, and A. Saldana. 2010. Human trypanosomiasis in the eastern region of the Panama Province: new endemic areas for Chagas disease. The American Journal of Tropical Medicine and Hygiene 82:580-582.

Campbell-Lendrum, D., L. Manga, M. Bagayoko, and J. Sommerfeld. 2015. Climate change and vector-borne diseases: what are the implications for public health research and policy? Philosophical transactions of the Royal Society of London.Series B, Biological sciences 370:10.1098/rstb.2013.0552.

Canals, M., L. Cruzat, M. C. Molina, A. Ferreira, and P. E. Cattan. 2001. Blood host sources of Mepraia spinolai (Heteroptera: Reduviidae), wild vector of chagas disease in Chile. Journal of medical entomology 38:303-307.

Cantey, P. T., S. L. Stramer, R. L. Townsend, H. Kamel, K. Ofafa, C. W. Todd, M. Currier, S. Hand, W. Varnado, E. Dotson, C. Hall, P. L. Jett, and S. P. Montgomery. 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 52:1922-1930.

Castro, J. A., and E. G. Diaz de Toranzo. 1988. Toxic effects of nifurtimox and benznidazole, two drugs used against American trypanosomiasis (Chagas' disease). Biomedical and environmental sciences : BES 1:19-33.

Cecere, M. C., G. M. Vazquez-Prokopec, R. E. Gürtler, and U. Kitron. 2006. Reinfestation sources for Chagas disease vector, Triatoma infestans, Argentina. Emerging Infectious Diseases 12:1096.

Cohen, J. E., and R. E. Gürtler. 2001. Modeling household transmission of American trypanosomiasis. Science (New York, N.Y.) 293:694-698.

Collini, M., H. Hauffe, S. Masseglia, F. Albonico, D. Arnoldi, X. Bailly, E. Bard, M. Galan, C. Rossi, and V. Tagliapietra. 2015. Identifying the last bloodmeal of questing wood tick nymphs (Ixodes ricinus L.) by DNA amplification: three approaches tested. Page Poster 2.18 in GERI 2015: Genes, Ecosystems and Risk of Infection.

Coura, J. R., and S. L. De Castro. 2002. A critical review on Chagas disease chemotherapy. Memórias do Instituto Oswaldo Cruz 97:3-24.

De Urioste-Stone, S. M., P. M. Pennington, E. Pellecer, T. M. Aguilar, G. Samayoa, H. D. Perdomo, H. Enriquez, and J. G. Juárez. 2015. Development of a community-based intervention for the control of Chagas disease based on peridomestic animal management: an eco-bio-social perspective. Transactions of the Royal Society of Tropical Medicine and Hygiene 109:159-167.

Dias, J. C. P. 2007. Southern Cone Initiative for the elimination of domestic populations of Triatoma infestans and the interruption of transfusion Chagas disease: historical aspects, present situation, and perspectives. Memórias do Instituto Oswaldo Cruz 102:11-18.

Dumonteil, E., and C. Herrera. 2017. Ten years of Chagas disease research: Looking back to achievements, looking ahead to challenges. PLoS neglected tropical diseases 11:e0005422.

Dumonteil, E., M.-J. Ramirez-Sierra, S. Pérez-Carrillo, C. Teh-Poot, C. Herrera, S. Gourbière, and E. Waleckx. 2018. Detailed ecological associations of triatomines revealed by metabarcoding and next-generation sequencing: implications for triatomine behavior and Trypanosoma cruzi transmission cycles. Scientific reports 8:4140.

Edwards, M. S., K. K. Stimpert, and S. P. Montgomery. 2017. Addressing the challenges of Chagas Disease: An emerging health concern in the United States. Infectious Diseases in Clinical Practice 25:118-125.

Elias, J. E., and S. P. Gygi. 2010. Target-decoy search strategy for mass spectrometry-based proteomics. Proteome Bioinformatics:55-71.

Feliciangeli, M. D., J. Rabinovich, and E. Fernandez. 1980. Resistencia al ayuno em Triatominos (Hemiptera, Reduviidae) venezolanos. Rev Inst Med Trop Sao Paulo 22.

Freitas, S. P. C., E. S. Lorosa, D. C. S. Rodrigues, A. L. C. Freitas, and T. C. M. Gonçalves. 2005. Feeding patterns of Triatoma pseudomaculata in the state of Ceara, Brazil. Revista de Saúde Pública 39:27-32.

Georgieva, A. Y., E. R. L. Gordon, and C. Weirauch. 2017. Sylvatic host associations of Triatominae and implications for Chagas disease reservoirs: a review and new host records based on archival specimens. PeerJ 5:e3826.

Gerber, S. A., J. Rush, O. Stemman, M. W. Kirschner, and S. P. Gygi. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proceedings of the National Academy of Sciences of the United States of America 100:6940-6945.

Goloborodko, A. A., L. I. Levitsky, M. V. Ivanov, and M. V. Gorshkov. 2013. Pyteomics—a Python framework for exploratory data analysis and rapid software prototyping in proteomics. Journal of the American Society for Mass Spectrometry 24:301-304.

Gómez-Díaz, E., and J. Figuerola. 2010. New perspectives in tracing vector-borne interaction networks. Trends in parasitology 26:470-476.

Gonçalves, T. C. M., D. S. Rocha, and R. A. Cunha. 2000. Feeding patterns of Triatoma vitticeps in the State of Rio de Janeiro, Brazil. Revista de Saúde Pública 34:348-352.

Goswami, T., X. Li, A. M. Smith, E. M. Luderowski, J. J. Vincent, J. Rush, and B. A. Ballif. 2012. Comparative phosphoproteomic analysis of neonatal and adult murine brain. PROTEOMICS 12:2185-2189.

Gottdenker, N. L., L. F. Chaves, J. E. Calzada, A. Saldaña, and C. R. Carroll. 2012. Host Life History Strategy, Species Diversity, and Habitat Influence

Trypanosoma cruzi Vector Infection in Changing Landscapes. PLoS neglected tropical diseases 6:e1884.
Greenwalt, D. E., Y. S. Goreva, S. M. Siljestrom, T. Rose, and R. E. Harbach. 2013. Hemoglobin-derived porphyrins preserved in a Middle Eocene blood-engorged mosquito. Proceedings of the National Academy of Sciences of the United States of America 110:18496-18500.
Gurevitz, J. M., M. S. Gaspe, G. F. Enriquez, Y. M. Provecho, U. Kitron, and R. E. Gürtler. 2013. Intensified surveillance and insecticide-based control of the Chagas disease vector Triatoma infestans in the Argentinean Chaco. PLoS neglected tropical diseases 7:e2158.
Gürtler, R., M. Cecere, M. Lauricella, M. Cardinal, U. Kitron, and J. Cohen. 2007. Domestic dogs and cats as sources of Trypanosoma cruzi infection in rural northwestern Argentina. Parasitology 134:69-82.
Gürtler, R., M. Cécere, D. Rubel, R. Petersen, N. Schweigmann, M. Lauricella, M. Bujas, E. Segura, and C. Wisnivesky-Colli. 1991. Chagas disease in north-west Argentina: infected dogs as a risk factor for the domestic transmission of Trypanosoma cruzi. Transactions of the Royal Society of Tropical Medicine and Hygiene 85:741-745.
Gürtler, R. E., and M. Cardinal. 2015. Reservoir host competence and the role of domestic and commensal hosts in the transmission of Trypanosoma cruzi. Acta Tropica 151:32-50.
Gürtler, R. E., M. C. Cecere, M. B. Castanera, D. Canale, M. A. Lauricella, R. Chuit, J. E. Cohen, and E. L. Segura. 1996a. Probability of infection with Trypanosoma cruzi of the vector Triatoma infestans fed on infected humans and dogs in northwest Argentina. American Journal of Tropical Medicine and Hygiene 55:24-31.
Gürtler, R. E., M. C. Cecere, D. P. Vazquez, R. Chuit, and J. E. Cohen. 1996b. Host-Feeding Patterns of Domiciliary Triatoma infestans (Hemiptera: Reduviidae) in Northwest Argentina: Seasonal and Instar Variation. Journal of medical entomology 33:15-26.
Gürtler, R. E., J. E. Cohen, M. C. Cecere, M. A. Lauricella, R. Chuit, and E. L. Segura. 1998. Influence of humans and domestic animals on the household prevalence of Trypanosoma cruzi in Triatoma infestans populations in northwest Argentina. The American Journal of Tropical Medicine and Hygiene 58:748-758.
Gürtler , R. E., and Z. E. Yadon. 2015. Eco-bio-social research on community-based approaches for Chagas disease vector control in Latin America. Transactions of the Royal Society of Tropical Medicine and Hygiene 109:91-98.
Hardison, R. C. 2012. Evolution of hemoglobin and its genes. Cold Spring Harbor perspectives in medicine 2:a011627.
Hashimoto, K., C. Cordon-Rosales, R. Trampe, and M. Kawabata. 2006. Impact of single and multiple residual sprayings of pyrethroid insecticides against Triatoma dimidiata (Reduviiade; Triatominae), the principal vector of Chagas disease in Jutiapa, Guatemala. American Journal of Tropical Medicine and Hygiene 75:226.

Henao-Martínez, A. F., K. Colborn, and G. Parra-Henao. 2017. Overcoming research barriers in Chagas disease—designing effective implementation science. Parasitology research 116:35-44.

Hotez, P. J., M. Alvarado, M.-G. Basáñez, I. Bolliger, R. Bourne, M. Boussinesq, S. J. Brooker, A. S. Brown, G. Buckle, and C. M. Budke. 2014. The global burden of disease study 2010: interpretation and implications for the neglected tropical diseases. PLoS neglected tropical diseases 8:e2865.

Hotez, P. J., E. Dumonteil, M. B. Cravioto, M. E. Bottazzi, R. Tapia-Conyer, S. Meymandi, U. Karunakara, I. Ribeiro, R. M. Cohen, and B. Pecoul. 2013. An unfolding tragedy of Chagas disease in North America. PLOS Negl Trop Dis 7:e2300.

Humair, P. F., V. Douet, F. Moran Cadenas, L. M. Schouls, I. Van De Pol, and L. Gern. 2007. Molecular identification of bloodmeal source in Ixodes ricinus ticks using 12S rDNA as a genetic marker. Journal of medical entomology 44:869-880.

Johnson, R. S., S. A. Martin, K. Biemann, J. T. Stults, and J. T. Watson. 1987. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Analytical chemistry 59:2621-2625.

Justi, S. A., and C. Galvão. 2017. The Evolutionary Origin of Diversity in Chagas Disease Vectors. Trends in parasitology 33:42-52.

Karlsson, A. O., and G. Holmlund. 2007. Identification of mammal species using species-specific DNA pyrosequencing. Forensic science international 173:16-20.

Keller, J. I., B. A. Ballif, R. M. S. Clair, J. J. Vincent, M. C. Monroy, and L. Stevens. 2017. Chagas disease vector blood meal sources identified by protein mass spectrometry. PLoS One 12:e0189647.

Keller, J. I., J. O. Schmidt, A. M. Schmoker, B. A. Ballif, and L. Stevens. 2018. Protein mass spectrometry extends temporal blood meal detection over polymerase chain reaction in mouse-fed Chagas disease vectors. Memórias do Instituto Oswaldo Cruz 113.

Kent, R. J. 2009. Molecular methods for arthropod bloodmeal identification and applications to ecological and vector-borne disease studies. Molecular ecology resources 9:4-18.

Kieran, T. J., N. L. Gottdenker, C. P. Varian, A. Saldaña, N. Means, D. Owens, J. E. Calzada, and T. C. Glenn. 2017. Blood meal source characterization using Illumina Sequencing in the Chagas Disease vector Rhodnius pallescens (Hemiptera: Reduviidae) in Panamá. Journal of medical entomology 54:1786-1789.

Kitano, T., K. Umetsu, W. Tian, and M. Osawa. 2007. Two universal primer sets for species identification among vertebrates. International journal of legal medicine 121:423-427.

Kjos, S. A., P. L. Marcet, M. J. Yabsley, U. Kitron, K. F. Snowden, K. S. Logan, J. C. Barnes, and E. M. Dotson. 2013. Identification of bloodmeal sources and Trypanosoma cruzi infection in triatomine bugs (Hemiptera: Reduviidae) from

residential settings in Texas, the United States. Journal of medical entomology 50:1126-1139.
Klotz, S. A., J. O. Schmidt, P. L. Dorn, C. Ivanyi, K. R. Sullivan, and L. Stevens. 2014. Free-roaming kissing bugs, vectors of Chagas disease, feed often on humans in the Southwest. The American Journal of Medicine 127:421-426.
Lardeux, F., C. Aliaga, and S. Depickère. 2016. Bias due to methods of parasite detection when estimating prevalence of infection of Triatoma infestans by Trypanosoma cruzi. Journal of Vector Ecology 41:285-291.
Laskay, Ü. A., L. Breci, I.-M. E. Vilcins, G. Dietrich, A. G. Barbour, J. Piesman, and V. H. Wysocki. 2013. Survival of host blood proteins in Ixodes scapularis (Acari: Ixodidae) ticks: a time course study. Journal of medical entomology 50:1282-1290.
Laskay, Ü. A., J. Burg, E. J. Kaleta, I.-M. E. Vilcins, S. R. Telford Iii, A. G. Barbour, and V. H. Wysocki. 2012a. Development of a host blood meal database: de novo sequencing of hemoglobin from nine small mammals using mass spectrometry.
Laskay, Ü. A., J. Burg, E. J. Kaleta, I. M. Vilcins, S. R. Telford Iii, A. G. Barbour, and V. H. Wysocki. 2012b. Development of a host blood meal database: de novo sequencing of hemoglobin from nine small mammals using mass spectrometry. Biolical Chemistry 393:195-201.
Lehane, M. J. 2005. The biology of blood-sucking in insects. Cambridge University Press.
Lian, J. B., G. S. Stein, A. J. Van Wijnen, J. L. Stein, M. Q. Hassan, T. Gaur, and Y. Zhang. 2012. MicroRNA control of bone formation and homeostasis. Nature Reviews Endocrinology 8:212-227.
Lima-Cordón, R. A., L. Stevens, E. S. Ortíz, G. A. Rodas, V. Abrego, C. Zúñiga, and C. Monroy. 2018. Implementation science: Epidemiology and feeding profiles of the Chagas vector Triatoma dimidiata prior to Ecohealth intervention for three locations in Central America. PLOS Negl Trop Dis accepted
Logue, K., J. B. Keven, M. V. Cannon, L. Reimer, P. Siba, E. D. Walker, P. A. Zimmerman, and D. Serre. 2016. Unbiased characterization of Anopheles mosquito blood meals by targeted high-throughput sequencing. PLoS neglected tropical diseases 10:e0004512.
Lucero, D. E., L. A. Morrissey, D. M. Rizzo, A. Rodas, R. Garnica, L. Stevens, D. M. Bustamante, and M. C. Monroy. 2013. Ecohealth interventions limit Triatomine reinfestation following insecticide spraying in La Brea, Guatemala. The American Journal of Tropical Medicine and Hygiene 88:630-637.
Lucero, D. E., W. Ribera, J. C. Pizarro, C. Plaza, L. W. Gordon, R. Peña Jr, L. A. Morrissey, D. M. Rizzo, and L. Stevens. 2014a. Sources of Blood Meals of Sylvatic Triatoma guasayana near Zurima, Bolivia, Assayed with qPCR and 12S Cloning. PLOS Negl Trop Dis 8:e3365.
Lucero, D. E., W. Ribera, J. C. Pizarro, C. Plaza, L. W. Gordon, R. Pena, Jr., L. A. Morrissey, D. M. Rizzo, and L. Stevens. 2014b. Sources of Blood Meals of Sylvatic Triatoma guasayana near Zurima, Bolivia, Assayed with qPCR and 12S Cloning. PLOS Negl Trop Dis 8:e3365.

Lundgren, D. H., S.-I. Hwang, L. Wu, and D. K. Han. 2010. Role of spectral counting in quantitative proteomics. Expert review of proteomics 7:39-53.

Luz, C., J. Fargues, and C. Romana. 2003. Influence of starvation and blood meal-induced moult on the susceptibility of nymphs of Rhodnius prolixus Stål (Hem., Triatominae) to Beauveria bassiana (Bals.) Vuill. infection. Journal of applied entomology 127:153-156.

Martínez-de la Puente, J., S. Ruiz, R. Soriguer, and J. Figuerola. 2013. Effect of blood meal digestion and DNA extraction protocol on the success of blood meal source determination in the malaria vector Anopheles atroparvus. Malaria journal 12:109.

Martínez-Ibarra, J. A., Y. Grant-Guillén, B. Nogueda-Torres, and F. Trujillo-Contreras. 2004. Influence of the blood meal source on the biology of Meccus longipennis (Hemiptera: Reduviidae) under laboratory conditions. Journal of the American Mosquito Control Association 20:328-330.

Melton, T., and C. Holland. 2007. Routine forensic use of the mitochondrial 12S ribosomal RNA gene for species identification. Journal of forensic sciences 52:1305-1307.

Mesquita, R. D., R. J. Vionette-Amaral, C. Lowenberger, R. Rivera-Pomar, F. A. Monteiro, P. Minx, J. Spieth, A. B. Carvalho, F. Panzera, D. Lawson, A. Q. Torres, J. M. C. Ribeiro, M. H. F. Sorgine, R. M. Waterhouse, M. J. Montague, F. Abad-Franch, M. Alves-Bezerra, L. R. Amaral, H. M. Araujo, R. N. Araujo, L. Aravind, G. C. Atella, P. Azambuja, M. Berni, P. R. Bittencourt-Cunha, G. R. C. Braz, G. Calderón-Fernández, C. M. A. Cararetо, M. B. Christensen, I. R. Costa, S. G. Costa, M. Dansa, C. R. O. Daumas-Filho, I. F. De-Paula, F. A. Dias, G. Dimopoulos, S. J. Emrich, N. Esponda-Behrens, P. Fampa, R. D. Fernandez-Medina, R. N. da Fonseca, M. Fontenele, C. Fronick, L. A. Fulton, A. C. Gandara, E. S. Garcia, F. A. Genta, G. I. Giraldo-Calderón, B. Gomes, K. C. Gondim, A. Granzotto, A. A. Guarneri, R. Guigó, M. Harry, D. S. T. Hughes, W. Jablonka, E. Jacquin-Joly, M. P. Juárez, L. B. Koerich, A. B. Lange, J. M. Latorre-Estivalis, A. Lavore, G. G. Lawrence, C. Lazoski, C. R. Lazzari, R. R. Lopes, M. G. Lorenzo, M. D. Lugon, D. Majerowicz, P. L. Marcet, M. Mariotti, H. Masuda, K. Megy, A. C. A. Melo, F. Missirlis, T. Mota, F. G. Noriega, M. Nouzova, R. D. Nunes, R. L. L. Oliveira, G. Oliveira-Silveira, S. Ons, I. Orchard, L. Pagola, G. O. Paiva-Silva, A. Pascual, M. G. Pavan, N. Pedrini, A. A. Peixoto, M. H. Pereira, A. Pike, C. Polycarpo, F. Prosdocimi, R. Ribeiro-Rodrigues, H. M. Robertson, A. P. Salerno, D. Salmon, D. Santesmasses, R. Schama, E. S. Seabra-Junior, L. Silva-Cardoso, M. A. C. Silva-Neto, M. Souza-Gomes, M. Sterkel, M. L. Taracena, M. Tojo, Z. J. Tu, J. M. C. Tubio, R. Ursic-Bedoya, T. M. Venancio, A. B. Walter-Nuno, D. Wilson, W. C. Warren, R. K. Wilson, E. Huebner, E. M. Dotson, and P. L. Oliveira. 2015. Genome of Rhodnius prolixus, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. Proceedings of the National Academy of Sciences 112:14936-14941.

Molaei, G., A. Farajollahi, J. J. Scott, R. Gaugler, and T. G. Andreadis. 2009. Human bloodfeeding by the recently introduced mosquito, Aedes japonicus japonicus,

and public health implications. Journal of the American Mosquito Control Association 25:210-214.
Monroy, C., D. M. Bustamante, S. Pineda, A. Rodas, X. Castro, V. Ayala, J. Quiñónes, and B. Moguel. 2009. House improvements and community participation in the control of Triatoma dimidiata re-infestation in Jutiapa, Guatemala. Cadernos de Saúde Pública 25:S168-S178.
Mota, J., J. C. Chacon, A. E. Gutiérrez-Cabrera, V. Sanchez-Cordero, R. A. Wirtz, R. Ordonez, F. Panzera, and J. M. Ramsey. 2007. Identification of blood meal source and infection with Trypanosoma cruzi of Chagas disease vectors using a multiplex cytochrome b polymerase chain reaction assay. Vector-Borne and Zoonotic Diseases 7:617-628.
Mudhasani, R., Z. Zhu, G. Hutvagner, C. M. Eischen, S. Lyle, L. L. Hall, J. B. Lawrence, A. N. Imbalzano, and S. N. Jones. 2008. Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells. The Journal of cell biology 181:1055-1063.
Nakagawa, J., C. Cordón-Rosales, J. Juárez, C. Itzep, and T. Nonami. 2003. Impact of residual spraying on Rhodnius prolixus and Triatoma dimidiata in the department of Zacapa in Guatemala. Memórias do Instituto Oswaldo Cruz 98:277-282.
Natarajan, C., F. G. Hoffmann, H. C. Lanier, C. J. Wolf, Z. A. Cheviron, M. L. Spangler, R. E. Weber, A. Fago, and J. F. Storz. 2015. Intraspecific polymorphism, interspecific divergence, and the origins of function-altering mutations in deer mouse hemoglobin. Molecular biology and evolution 32:978-997.
National Research Council. 2010. Guide for the care and use of laboratory animals. National Academies Press.
Nattero, J., G. Leonhard, C. S. Rodríguez, and L. Crocco. 2011. Influence of the quality and quantity of blood ingested on reproductive parameters and life-span in Triatoma infestans (Klug). Acta Tropica 119:183-187.
Ngo, K. A., and L. D. Kramer. 2003. Identification of mosquito bloodmeals using polymerase chain reaction (PCR) with order-specific primers. Journal of medical entomology 40:215-222.
Niare, S., J.-M. Berenger, C. Dieme, O. Doumbo, D. Raoult, P. Parola, and L. Almeras. 2016. Identification of blood meal sources in the main African malaria mosquito vector by MALDI-TOF MS. Malaria journal 15:1.
Noriega, F. G. 1992. Autogeny in three species of Triatominae: Rhodnius prolixus, Triatoma rubrovaria, and Triatoma infestans (Hemiptera: Reduviidae). Journal of medical entomology 29:273-277.
Önder, Ö., W. Shao, B. D. Kemps, H. Lam, and D. Brisson. 2013. Identifying sources of tick blood meals using unidentified tandem mass spectral libraries. Nature communications 4:1746.
Önder, Ö., W. Shao, H. Lam, and D. Brisson. 2014. Tracking the sources of blood meals of parasitic arthropods using shotgun proteomics and unidentified tandem mass spectral libraries. Nature protocols 9:842-850.

Orantes, L. C. 2017. Assessing Community Dynamics and Colonization Patterns of Triatoma dimidiata and Other Biotic Factors Associated with Chagas Disease Prevalence in Central America. The University of Vermont and State Agricultural College.

Orantes, L. C., C. Monroy, P. L. Dorn, L. Stevens, D. M. Rizzo, L. Morrissey, J. P. Hanley, A. G. Rodas, B. Richards, and K. F. Wallin. 2018. Uncovering vector, parasite, blood meal and microbiome patterns from mixed-DNA specimens of the Chagas disease vector Triatoma dimidiata. PLoS neglected tropical diseases 12:e0006730.

Padrón, G., and G. B. Domont. 2014. Two decades of proteomics in Latin America: a personal view. Journal of proteomics 107:83-92.

Pellecer, M. J., P. L. Dorn, D. M. Bustamante, A. Rodas, and M. C. Monroy. 2013. Vector blood meals are an early indicator of the effectiveness of the Ecohealth approach in halting Chagas transmission in Guatemala. The American Journal of Tropical Medicine and Hygiene 88:638-644.

Peña, V. H., G. J. Fernández, A. M. Gómez-Palacio, A. M. Mejía-Jaramillo, O. Cantillo, and O. Triana-Chávez. 2012. High-resolution melting (HRM) of the cytochrome B gene: a powerful approach to identify blood-meal sources in Chagas disease Vectors. PLOS Negl Trop Dis 6:e1530.

Pineda, V., A. Saldaña, I. Monfante, A. Santamaría, N. Gottdenker, M. Yabsley, G. Rapoport, and J. Calzada. 2011. Prevalence of trypanosome infections in dogs from Chagas disease endemic regions in Panama, Central America. Veterinary parasitology 178:360-363.

Pinto, J., D. M. Roellig, R. H. Gilman, M. Calderón, C. Bartra, R. Salazar, C. Bern, J. Ancca-Juárez, M. Levy, C. Náquira, and V. Cama. 2012. Temporal differences in blood meal detection from the midguts of Triatoma infestans. Revista do Instituto de Medicina Tropical de São Paulo 54:83-88.

Pizarro, J. C., D. Lucero, and L. Stevens. 2007. A method for the identification of guinea pig blood meal in the Chagas disease vector, Triatoma infestans. Kinetoplastid biology and disease 6:1.

Pizarro, J. C., and L. Stevens. 2008. A new method for forensic DNA analysis of the blood meal in Chagas disease vectors demonstrated using Triatoma infestans from Chuquisaca, Bolivia. PLoS One 3:e3585.

Post, R., P. Flook, and A. Millest. 1993. Methods for the preservation of insects for DNA studies. Biochemical systematics and ecology 21:85-92.

Rabinovich, J. E., U. D. Kitron, Y. Obed, M. Yoshioka, N. Gottdenker, and L. F. Chaves. 2011. Ecological patterns of blood-feeding by kissing-bugs (Hemiptera: Reduviidae: Triatominae). Memórias do Instituto Oswaldo Cruz 106:479-494.

Ramírez, J. D., B. Turriago, G. Tapia-Calle, and F. Guhl. 2013. Understanding the role of dogs (Canis lupus familiaris) in the transmission dynamics of Trypanosoma cruzi genotypes in Colombia. Veterinary parasitology 196:216-219.

Rassi, A., and J. A. Marin-Neto. 2010. Chagas disease. The Lancet 375:1388-1402.

Rizzo, N. R., B. A. Arana, A. Diaz, C. Cordon-Rosales, R. E. Klein, and M. R. Powell. 2003. Seroprevalence of Trypanosoma cruzi infection among school-age children

in the endemic area of Guatemala. The American Journal of Tropical Medicine and Hygiene 68:678-682.
Sánchez-Valdéz, F. J., C. Pérez Brandán, A. Ferreira, and M. Á. Basombrío. 2015. Gene-deleted live-attenuated Trypanosoma cruzi parasites as vaccines to protect against Chagas disease. Expert review of vaccines 14:681-697.
Sasaki, H., R. Rosales, and Y. Tabaru. 2003. Host feeding profiles of Rhodnius prolixus and Triatoma dimidiata in Guatemala (Hemiptera: Reduviidae: Triatominae). Medical Entomology and Zoology 54:283-289.
Sayers, E. W., T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio, R. Edgar, S. Federhen, M. Feolo, L. Y. Geer, W. Helmberg, Y. Kapustin, D. Landsman, D. J. Lipman, T. L. Madden, D. R. Maglott, V. Miller, I. Mizrachi, J. Ostell, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, M. Shumway, K. Sirotkin, A. Souvorov, G. Starchenko, T. A. Tatusova, L. Wagner, E. Yaschenko, and J. Ye. 2009. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 37:D5-15.
Song, Y., Ü. A. Laskay, I.-M. E. Vilcins, A. G. Barbour, and V. H. Wysocki. 2015. Top-Down-Assisted Bottom-Up Method for Homologous Protein Sequencing: Hemoglobin from 33 Bird Species. Journal of the American Society for Mass Spectrometry 26:1875-1884.
Souza, R. d. C. M. d., A. C. Soares, C. L. Alves, E. S. Lorosa, M. H. Pereira, and L. Diotaiuti. 2011. Feeding behavior of Triatoma vitticeps (Reduviidae: Triatominae) in the state of Minas Gerais, Brazil. Memórias do Instituto Oswaldo Cruz 106:16-22.
Steen, H., and M. Mann. 2004. The ABC's (and XYZ's) of peptide sequencing. Nature reviews Molecular cell biology 5:699-711.
Stevens, L., P. L. Dorn, J. Hobson, N. M. de la Rua, D. E. Lucero, J. H. Klotz, J. O. Schmidt, and S. A. Klotz. 2012. Vector blood meals and Chagas disease transmission potential, United States. Emerging Infectious Diseases 18:646-649.
Stevens, L., M. C. Monroy, A. G. Rodas, and P. L. Dorn. 2014. Hunting, Swimming, and Worshiping: Human Cultural Practices Illuminate the Blood Meal Sources of Cave Dwelling Chagas Vectors (Triatoma dimidiata) in Guatemala and Belize. PLOS Negl Trop Dis 8:e3047.
Stimpert, K. K., and S. P. Montgomery. 2010. Physician awareness of Chagas disease, USA. Emerging Infectious Diseases 16:871.
Storz, J. F., A. M. Runck, S. J. Sabatino, J. K. Kelly, N. Ferrand, H. Moriyama, R. E. Weber, and A. Fago. 2009. Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. Proceedings of the National Academy of Sciences 106:14450-14455.
Sutherst, R. W. 2004. Global change and human vulnerability to vector-borne diseases. Clinical microbiology reviews 17:136-173.
Vasquez, A. M. d., F. E. Samudio, A. Saldaña, H. M. Paz, and J. E. Calzada. 2004. Eco-epidemiological aspects of Trypanosoma cruzi, Trypanosoma rangeli and their vector (Rhodnius pallescens) in Panama. Revista do Instituto de Medicina Tropical de São Paulo 46:217-222.

Vazquez-Prokopec, G. M., C. Spillmann, M. Zaidenberg, U. Kitron, and R. E. Gürtler. 2009. Cost-effectiveness of Chagas disease vector control strategies in northwestern Argentina. PLoS neglected tropical diseases 3:e363.

Waleckx, E., J. Camara-Mejia, M. J. Ramirez-Sierra, V. Cruz-Chan, M. Rosado-Vallado, S. Vazquez-Narvaez, R. Najera-Vazquez, S. Gourbière, and E. Dumonteil. 2015. An innovative ecohealth intervention for Chagas disease vector control in Yucatan, Mexico. Transactions of the Royal Society of Tropical Medicine and Hygiene 109:143-149.

Waleckx, E., J. Suarez, B. Richards, and P. L. Dorn. 2014. Triatoma sanguisuga blood meals and potential for Chagas disease, Louisiana, USA. Emerg Infect Dis 20:2141-2143.

Walker, J. A., D. A. Hughes, B. A. Anders, J. Shewale, S. K. Sinha, and M. A. Batzer. 2003. Quantitative intra-short interspersed element PCR for species-specific DNA identification. Analytical Biochemistry 316:259-269.

Walker, J. A., D. A. Hughes, D. J. Hedges, B. A. Anders, M. E. Laborde, J. Shewale, S. K. Sinha, and M. A. Batzer. 2004. Quantitative PCR for DNA identification based on genome-specific interspersed repetitive elements. Genomics 83:518-527.

Wickramasekara, S., J. Bunikis, V. Wysocki, and A. G. Barbour. 2008. Identification of residual blood proteins in ticks by mass spectrometry proteomics. Emerging Infectious Diseases 14.

World Health Organization. 2012. Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. World Health Organ Tech Rep Ser:v-xii, 1-100.

World Health Organization. 2014. A global brief on vector-borne diseases.

World Health Organization. 2015. Chagas disease in Latin America: an epidemiological update based on 2010 estimates. Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire 90:33-44.

World Health Organization. 2016. Chagas disease (American trypanosomiasis). Fact sheet N 340 (2015). World Health Organization.

World Health Organization. 2017. Vector borne disease World Health Organization, Geneva, Switzerland.

World Health Organization, and UNICEF. 2017. Global vector control response 2017-2030.

Yang, H., and R. A. Zubarev. 2010. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. ELECTROPHORESIS 31:1764-1772.

Yoshioka, K., J. Nakamura, B. Pérez, D. Tercero, L. Pérez, and Y. Tabaru. 2015. Effectiveness of large-scale Chagas disease vector control program in Nicaragua by residual insecticide spraying against Triatoma dimidiata. The American Journal of Tropical Medicine and Hygiene 93:1231-1239.

Yssouf, A., P. Parola, A. Lindström, T. Lilja, G. L'Ambert, U. Bondesson, J.-M. Berenger, D. Raoult, and L. Almeras. 2014. Identification of European mosquito species by MALDI-TOF MS. Parasitology research 113:2375-2378.

Zeledón, R., V. M. Guardia, A. Zuñiga, and J. C. Swartzwelder. 1970. Biology and ethology of Triatoma Dimidiata (Latreille, 1811) II. Life span of adults and fecundity and fertility of females. Journal of medical entomology 7:462-469.
Zeledon, R., and J. E. Rabinovich. 1981. Chagas disease: an ecological appraisal with special emphasis on its insect vectors. Annual Review of Entomology 26:101-133.