University of Vermont
UVM ScholarWorks

UVM College of Arts and Sciences College Honors Theses

Undergraduate Theses

2017

Protein Identification and Analysis of Blood Meal Digestion in Different Chagas Disease Insect Vector Body Parts Over Time

Emily M. Eakin

Follow this and additional works at: https://scholarworks.uvm.edu/castheses

Recommended Citation

Eakin, Emily M., "Protein Identification and Analysis of Blood Meal Digestion in Different Chagas Disease Insect Vector Body Parts Over Time" (2017). *UVM College of Arts and Sciences College Honors Theses*. 98.
https://scholarworks.uvm.edu/castheses/98

This Undergraduate Thesis is brought to you for free and open access by the Undergraduate Theses at UVM ScholarWorks. It has been accepted for inclusion in UVM College of Arts and Sciences College Honors Theses by an authorized administrator of UVM ScholarWorks. For more information, please contact donna.omalley@uvm.edu.

Protein Identification and Analysis of Blood Meal Digestion in Different Chagas Disease Insect Vector Body Parts Over Time

By Emily Eakin

Abstract

Chagas disease is a potentially life-threatening blood-borne infection that is transmitted to vertebrate hosts by the parasite *Trypanosoma cruzi* (Kinetoplastea: Tripanosomatida), which is carried by hematophagous Triatominae (Hemiptera: Reduviidae) insect vector species that deposit infected feces when taking a vertebrate blood meal. There is currently no accepted treatment for the chronic stages of Chagas disease, posing significant challenges since clinical manifestations and diagnosis may not occur until decades after initial infection. With the goal of lowering overall human infection rate, it is important to identify the pool of vertebrate blood meal sources available to insect vectors to design targeted ecohealth disease prevention measures. As a step in developing methodologies to identify these host species, I analyzed the blood meal composition of experimentally-fed Triatominae insect vectors immediately after and 8 weeks after a blood meal to evaluate: (1) the passage of the blood meal through progressive body parts (head, thorax, upper and lower abdomen) at the two time points and (2) the relative efficiency of identifying two blood proteins (albumin and hemoglobin) at the two time points.

This control experiment focuses on liquid chromatography tandem mass spectrometry (LC-MS/MS) identification of two blood proteins, hemoglobin and albumin, within successive segments of the insect digestive tract to estimate the protein detectable in the insect at two time points. *Triatoma protracta* experimentally-fed on

house mouse (*Mus musculus)* were collected immediately after feeding or 8 weeks post-feeding (with no subsequent feeding). Body parts of the insects were compared with respect to the hemoglobin and albumin protein peptides present. Hemoglobin and albumin were detected in all body parts immediately after a blood meal. In addition, the high number of spectral counts matching the expected mouse blood source confirms the efficacy of LC-MS/MS identification. By 8 weeks post-blood meal, we were still able to detect hemoglobin and albumin in several insect parts, but with lower spectral count quantities. I conclude that elucidating timelines for the degradation of blood proteins under controlled conditions could change the future of blood source analysis by allowing more precise dating of previous blood meals, which is important information for developing ecohealth strategies.

Introduction

Chagas disease is a potentially life-threatening, blood-borne infection that is caused by the parasite *Trypanosoma cruzi*. This parasite is most often transmitted between mammalian hosts by Triatominae insect vectors, which become infected after ingestion of blood containing trypomastigotes, which are the infective form of *T. cruzi*. Disease transmission occurs when the insect deposits infected feces while taking a subsequent blood meal from another mammal (**Figure 1**) (cdc.gov, 2016). The onset of infection in the host characterizes the acute phase of Chagas disease, which is described

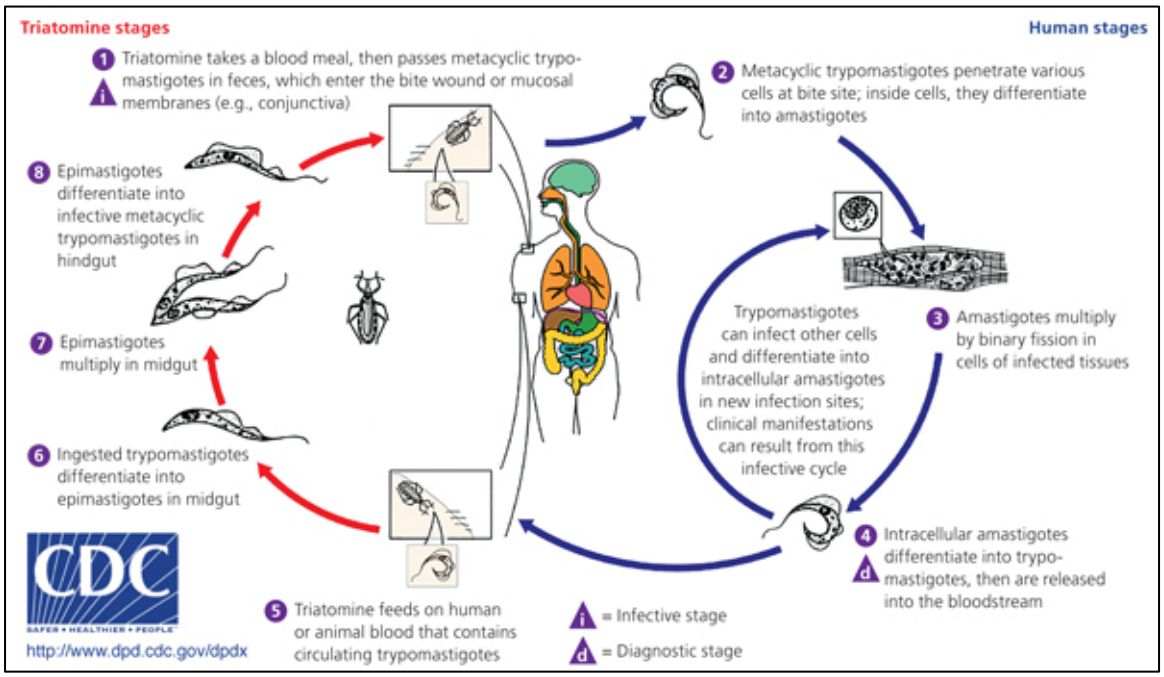


Figure 1. *This schematic of the* T. cruzi *parasite life cycle demonstrates the life stages while entering and infecting a human host (blue arrows) and while passing through an insect vector (red arrows). Triangular labels indicate points where* T. cruzi *is in the infective stage or diagnostic stage.*

as clinically separate from the chronic, latent disease phase (Kirchhoff, 2011). The biphasic nature of the infection contributes to the clinical difficulties that cause Chagas disease to be costly; as of 2000, the worldwide disease costs more than US$7-19 billion annually (Lee et al., 2013). With about 8 million people carrying the *T. cruzi* parasite and

more than 20,000 annual deaths attributed to Chagas disease, global public health research in this field has increased in recent years, yet there is still no vaccine (Lee et al. 2013), and drug treatment is controversial. Significant difficulties surrounding the diagnosis, treatment, and pediatric management of Chagas disease have refocused research into devising alternative prevention measures to lower the infection rates preemptively. Below, I will outline these three difficulties, then describe how my research of blood meal identification may provide an alternate approach to reduce the global burden of Chagas disease.

Firstly, while the acute disease phase is associated with some identifiable symptoms, infection often occurs undiagnosed. Acute symptoms include chagoma, a lesion at the site of parasite entry, and Romaña's sign, which presents as swelling of the eye, but only occurs in a small proportion of patients, and can be the result of other infections as well (Kirchhoff, 2011). Despite the potential for these symptoms, diagnoses in the acute phase are still uncommon, since some patients present with mild flu-like symptoms or symptoms too non-irritating to seek treatment. Clinically speaking, the chronic phase is of greater concern, since 10-30% of those infected will present with severe symptoms even decades after initial infection (Kirchhoff, 2011). Cardiac problems are the most common chronic manifestation, with symptoms ranging from mild inflammation and dysrhythmias to sudden death from complete ventricular blockage (Ozaki et al., 2011). Other chronic manifestations include megaesophagus and megacolon that lead to digestive complications and lesions in the gastrointestinal tract (Kirchhoff, 2011). Diagnosis in the chronic phase can be accomplished with specific antibodies and

potentially PCR, yet many of these diagnostic approaches have low efficacy (Brasil et al., 2010).

In addition to struggles surrounding accurate and timely diagnosis, other research efforts have attempted to identify an anti-parasitic drug treatment that is effective against *T. cruzi.* Only two chemotherapeutic drugs, Benznidazole and Nifurtimox are currently considered effective Chagas disease treatment drugs, but only if administered soon after the initial infection, since efficacy during the chronic phase is controversial (Saloman, 2011; Manarin et al., 2013 Boscardin et al., 2010). Both drugs have similar cure rates during the acute phase of about 70%, but are effective less than 10% of the time in the chronic phase (Kirchhoff, 2011). Additionally, both drugs have significant side effects that may include vomiting, anorexia, and neurologic symptoms ranging from restlessness to seizures (Saloman, 2011). These severe side effects are of particular concern, since Chagas disease is prevalent in areas with high rates of malnutrition, often rendering Benznidazole and Nifurtimox impracticable. Furthermore, even patients that seek treatment for the acute phase of the disease may be unlikely to continue an uncomfortable chemotherapy treatment for symptoms that alone seem tolerable (Kirchhoff, 2011). For the many patients who fail to be diagnosed during the acute phase, if they enter the chronic phase, there is no cure, and symptom management often includes surgery and is expensive and invasive (Kirchhoff, 2011).

Finally, another major concern is the disease management in pediatric patients. In addition to traditional vector transmission, transmission may occur from mother to fetus with up to 5% congenital transmission rate (Howard et al., 2014). Coupled with global migration, the potential 15,000 cases of this congenital transmission each year can

increase risk to non-endemic countries (Manarin et al., 2013). There is no clearly effective approach for blocking congenital transmission, so Chagas disease is treated more aggressively in children and in young women that may transmit the infection to their offspring. Clinical consensus appears to suggest that all infected children under 17 should be given Benznidazole or Nifurtimox chemotherapy, but the efficacy in protecting the fetuses of infected mothers is largely unknown (Sosa-Estani et al., 2012).

These aforementioned challenges associated with diagnosis, treatment, and pediatric disease management have helped spur research devoted to reducing the risk of vector transmission. In theory, a better understanding of Chagas disease host-vector interactions could be used to identify and manage environmental risk factors, reducing the global disease burden preemptively. To understand these host-vector relationships, we must determine the geographical variation in the species involved in the transmission cycle, and use the details of this interaction to direct ecohealth approaches to best fit the community.

To determine spatial and temporal variation in species interaction, we can investigate the contents of the insect gut and determine the source of the blood meal. Recent research indicates that hemoglobin-based mass-spectrometry (MS) and protein analysis may be a useful complementary method to existing DNA-based techniques for identifying blood meal composition of Triatominae insect vectors (Keller et al., 2017). Host blood protein specificity provides a unique opportunity to determine the type of vertebrate hosts that are likely to transmit disease to humans. Understanding common hosts in different at-risk populations could allow more efficient resource allocation to effective ecohelath approaches (Monroy et al., 2009).

Most mammals can act as a host for Triatomine insect vector species, and while non-mammal blood meals are possible, most epidemiologically relevant *T. cruzi* transmission occurs between domestic and anthropogenic mammals (Stevens et al., 2013). Domestic mammals are those that live alongside humans in their dwelling, such as cats and dogs. Anthropogenic mammals are those that live around human habitations, such as mice and rats.

Historically, agriculture has played an important role in the connection of humans to animal pathogens via anthropogenic mammals (Corondado et al., 2009). Human contact with farm mammals and other small anthropogenic animals that thrive on agricultural farmland increase the risk of infection by multiple pathways. Overlap between rural Central and South American communities coupled with increasing globalization ties populations with endemic Chagas disease to the rest of the world. In addition to caring for the 8 million infected in Latin America, to manage global disease risk, we must look back to small-scale *T. cruzi* parasite transmission by determining the blood meal sources of insect vector species.

Determining the blood meal sources of the insect vector comes with certain challenges and has proven difficult for insects stored for extended time periods (Waleckx et al., 2014, Stevens et a., 2012). As previous research has shown, applying mass spectrometry-based methods holds promise for blood meal identification in arthropod vectors (Laskay et al, 2012, 2013; Önder et al., 2013, 2014). Exploiting the taxonomic specificity of hemoglobin and albumin blood proteins allows precise identification of up to several recent blood meal sources, providing a multidimensional depiction of the species involved in the transmission of Chagas disease (Laskay et al, 2012, 2013; Önder

et al., 2013, 2014). Since hemoglobin and albumin are highly abundant in blood plasma, they are potentially good target proteins in a blood meal study (Laskay et al, 2012; 2013). In addition, these blood proteins are sufficiently conserved across species, with identifiable genetic variations in the proteins that allow for genus and species level taxonomic identification (Önder et al., 2013). Performing a timeline analysis by measuring the abundance of both albumin and hemoglobin in a sample could provide an estimate of the timing of prior blood meals based on residue protein within the insect vector gut.

The following three research aims intend to evaluate this novel timeline analysis while elucidating the ability to detect and identify multiple blood proteins throughout the digestive system of the insect:

First, traces of a blood meal can be found throughout the digestive system (Gürtler et al., 2014, Pizarro and Stevens 2008). I tested for the presence of host blood meal in each part of the insect vector digestive system (mouth-head, foregut-thorax, midgut-upper abdomen, hindgut-lower abdomen). Hypothesis 1 of this study is that hemoglobin from a vertebrate blood meal passes through the insect and traces of identifiable protein remain in the lower digestive tract. Insects have a complete digestive system, thus host blood meal must pass through all consecutive body parts, and may be found in low levels even in non-storage parts of the system (Gullan et al., 2014).

Second, I extend the time point and digestion parameters above to a second blood protein, albumin. Hemoglobin is often used because it is a very stable protein, but using a second protein may allow us to estimate when a blood meal was taken by the insect vector by comparing the rates of decay (Laskay et al., 2013; Wickramasekara et al.,

2008). The second hypothesis is that because hemoglobin is more stable than albumin, hemoglobin peptides will be identifiable in the insect gut for longer and/or in larger quantities than albumin, meaning more peptides will be identified in freshly fed vectors than vectors collected 8 weeks post feeding.

Hypothesis 3 is that there is a time threshold up to which the vertebrate blood meal source can be identified (Martínez-de la Puente et al., 2013; Oshaghi et al., 2006). I compared insects collected immediately after feeding with one collected 8 weeks after feeding. I hypothesized that if the insect gut degrades blood meal protein over time, then the ability to detect and identify peptides will decrease over time.

Methods

Feeding and sampling of insects:

This study used insects from a laboratory-reared colony of *Triatoma protracta* maintained at the Southwestern Biological Institute in Tucson, AZ. The Southwestern Biological Institute of Animal Care and Use Committee approved all procedures using vertebrate animals.

The three insects used for this study were part of a group collected after enclosing as adults and allowing a single blood meal feeding on *Mus musculus.* Within an hour of feeding, two freshly fed insects (hereafter referred to as 0 wk A and 0 wk B) were preserved in 95 % ethanol and 5 % glycerol. The third insect was maintained until

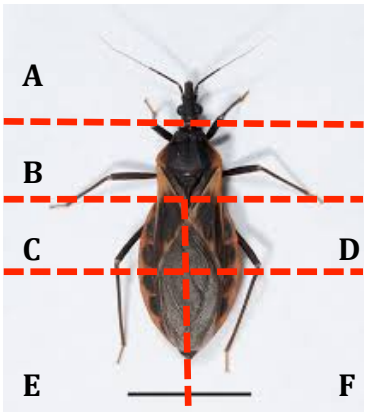


Figure 2 *Schematic of the Triatominae dissection. Sections A-C, and E were used for mass-spectrometry analysis. Portion A is the head, B is the thorax, C is the left upper abdomen (LUA) and E is the left lower abdomen (LLA).*

eight-weeks post-feeding (hereafter referred to as 8 wk) without any additional blood meals, and then preserved as described above. Within 1-4 weeks after collection, specimens were shipped to the University of Vermont (UVM) for subsequent dissection and evaluation. Upon receipt at UVM, insects were stored at -20°C until preparation for mass spectrometry. Insects were dissected to evaluate the blood proteins detectable in different body parts as shown in **Figure 2**.

SDS-PAGE, trypsin digestion, and mass spectrometry

Proteins were extracted from each insect body part using the extraction protocol from Keller et. al (2017, *in prep*). In short, denaturing sampling buffer is added to insect samples proportionally by mass, with 200 µl of buffer for every 100 mg of sample. Samples were ground and heated at 95°C for 5 minutes, and then

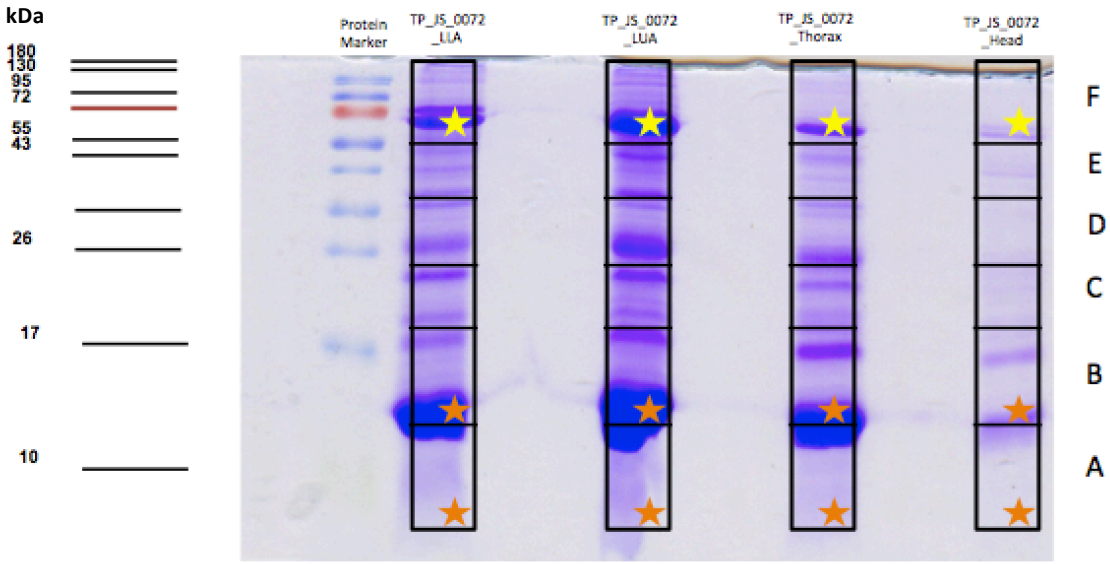


Figure 3 *SDS-PAGE of multiple body parts of a freshly fed Triatominae bug. The bug was sampled immediately after taking a blood meal and protein was present in each body part portion, from left to right; left lower abdomen (LLA), left upper abdomen (LUA), thorax, head. The band at around 16 kDa corresponds to hemoglobin, and the band at 65-70 kDa corresponds to albumin. The grid overlay in each lane serves as a template for cutting the gel, where orange starred boxes contain the hemoglobin and the yellow starred boxes contain the albumin.*

vortexed and centrifuged. Of this resultant extract, 20 µl samples were separated using 15% SDS-PAGE (37.5 acrylamide: 1 bis-acrylamide), and stained with Coomassie Blue, as shown in **Figure 3**. The gel was cut into 1mm cubes according to the grid overlay shown in **Figure 3** to isolate the regions including the proteins of interest, albumin and hemoglobin. *M. musculus* albumin (~608 amino acids) is around 65-70 kDa, and both alpha and beta subunits of hemoglobin are around 16 kDa (142 amino acids and 147 amino acids, respectively).

Gel pieces were prepared for mass spectrometry as previously described [Ballif et al. J Proteome Res. 2006 May; 5(9):2372-9]. In short, gel pieces were washed with 900 µl water, and then destained by incubating at 37 °C with 50% acetonitrile (MeCN) and 50 mM ammonium bicarbonate (NH_4HCO_3). Gel pieces were then immersed in 100% MeCN until fully dehydrated, then digested with 25 µl trypsin (6 ng/µl) and 50mM NH_4HCO_3 for 16 hours at a 37 °C incubation. The digestion solution was removed and collected and the gel pieces were further extracted with 100% MeCN. The MeCN extraction and digestion solutions were combined and peptides were dried via speed-vacuum. Tubes of extracted peptides were frozen at -80 °C until separation using high performance liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) using a LTQ-XL linear ion trap orbitrap mass spectrometer as described previously [Ballif et al. J Proteome Res. 2008; 7:311-8] except that all spectra were collected in the Orbitrap.

LC-MS/MS analysis and peptide identification

LC-MS/MS outputs spectral data that represent fragmented proteins as peptide masses and fragments thereof that can be accurately matched to sequences when referenced to a protein database. Separate custom forward and reverse concatenated databases were constructed for vertebrate hemoglobin and albumin using GenBank (Benson et al., 2007). The databases each consist of either all GenBank entries with "hemoglobin" or all GenBank entries with "albumin" in any search field as of 20 January 2016 and 26 October 2016, respectively (Benson et al., 2007). These custom databases were used by SEQUEST (Thermo Electron V26.12) when searching mass spectra using the target-decoy approach outlined by Elias and Gygi (2010) but modified to allow for the variable oxidation of methionine and acrylamidation of cysteine.

After matching spectra to databases, the following filtering criteria were applied: only peptides with XCorr values ≥2.5 (z=2) or 3 (z=3); observed precursor masses +/- 5 PPM of theoretical masses; unique ΔCn values ≥ 0.01; and no missed trypsin cleavages except at ragged ends (i.e., the 4 amino acids at the extreme N or C termini).

Because LC-MS/MS does not directly sequence a peptide, software is used to create a hypothetical spectrum of the masses associated with *in silico* trypsin-digested fragments of proteins [Pyteomics python tools 3.4.1] in the hemoglobin and albumin databases described above. Experimental spectral data are matched to these constructed databases (Goloborodko et al., 2013). Taxonomic lineage data were gathered from the NCBI taxonomy database for each vertebrate protein entry,

and parent protein taxonomic information was linked to peptides and stored in a relational database for subsequent retrieval (Sayers et al., 2009). Because a particular peptide often matches multiple taxa, multiple matches were summarized in pivot tables using Microsoft Excel (2011) for each sample (**Appendix A**). For all results, the sample size is too small for statistical analyses.

To manage taxonomic identification data for each insect and body part, tables were constructed with columns for each peptide and its spectral count and rows for the matching taxa (**Table 1**). Each pipeline includes counts of the spectral matches per taxa, an un-weighted measure of how many identified peptides match possible source species. The percent amino acids matching and the percent peptides matching were calculated for each taxon based on the number of amino acids per peptide, the number of peptides and

Table 1. Example pipeline summary of 0wk-B head hemoglobin(Hb). Amino acid sequences of all Hb peptides identified in this sample are listed from left to right in protein amino acid sequence order, adjacent to variants if they exist. Taxonomic affiliations are listed from class to species, and the number of species matching each peptide is highlighted and defined below if the number of species is lesss than 5. Red highlight=1 species, yellow =2 species, green=3 species, blue =4 species.

Sample: 0072	non-redundant peptides identified in sample																							
M. musculus NP_032244.2, BAG16710.1	VLSGEDKSNIK	IGGHGAEYGAEALER	MFASFPTTK	TYFPHFDVSHGSAQVK	VADALANAAGHLDDLPGAL	VADALATAAGHLDDLPGAL	VGDALGNAWAHLDDLPGAL	VAEALATAAGHLDDLPGAL	KVADALANAAGHLDDLPGI	LRVDPVNFK	FLASVSTVLTSK	VHLTDAEK	VNSDEVGGEALGR	NVADEVGGEALGR	VNADEVGGEALGR	YFDSFGDLSSASAIMGNAK	VITAFNDGLNHLDSLK	KVITAFNDGLNHLDSLK	LLGNMIVIVLGHHLGK	DFTPAAQAAFQK	VVAGVAAALAHK	VVAGVAAALAHKYH	
no. amino acids/peptide	11	15	9	16	29	29	29	29	30	9	12	8	13	13	13	19	16	17	16	12	12	14	Total 371
no. peptide variants	1	1	1	1	5					1	1	1	3			1	2		1	1	2		22
spectral count	2	2	2	6	1	2	4	1	2	3	23	2	3	2	1	113	2	6	5	2	4	3	191
taxanomic affiliations																							range
no. of classes	1	1	1	1	1	1	1	1	1	3	2	1	1	1	1	1	1	1	1	1	1	1	(1- 3)
no. of orders	1	1	1	2	1	1	1	1	2	51	9	2	1	1	1	1	1	1	1	2	2	2	(1 - 51)
no. of families	1	6	2	7	2	1	1	1	2	128	28	4	1	1	3	1	1	1	1	4	2	2	(1 - 128)
no. of genera	1	7	2	11	2	1	1	1	3	291	63	8	1	1	4	1	1	1	2	5	2	2	(1 - 291)
no. of species	1	9	2	19	3	2	1	1	3	443	94	12	3	1	6	8	4	4	9	13	4	4	(1 - 443)

Species reported with peptide																							Total peptide matches per taxa	Total peptide non-matches per taxa	Percent amino acids identified matching	Percent peptides identified matching	Percent spectral count matching
Mus musculus	1	1	1	1	1				1	1	1	1	1		1	1	1	1	1	1	1	1	18	4	73.0%	81.8%	95.29%
Mus spretus															1	1	1	1	1	1	1	1	8	14	32.1%	36.4%	71.20%
Mus spicilegus													1			1	1	1	1	1			6	16	25.1%	27.3%	68.59%
Mus macedonicus													1			1			1				3	19	12.9%	13.6%	63.35%
Mus cervicolor																1					1	1	3	19	12.1%	13.6%	62.83%
Nannospalax ehrenbergi		1		1	1				1	1	1	1								1			8	14	35.3%	36.4%	21.47%
Nannospalax galili		1		1	1				1	1	1									1			7	15	33.2%	31.8%	20.42%
Peromyscus maniculatus		1		1		1				1	1	1											6	16	24.0%	27.3%	19.90%
Cricetomys gambianus		1	1	1						1	1												5	17	16.4%	22.7%	18.85%
Peromyscus leucopus		1		1		1				1	1												5	17	21.8%	22.7%	18.85%
Tarsius syrichta										1	1										1	1	4	18	12.7%	18.2%	17.28%
Tamiasciurus hudsonicus								1		1	1			1									4	18	17.0%	18.2%	15.18%
Mus pahari																	1	1					2	20	8.9%	9.1%	4.19%
Chalinolobus morio							1			1													2	20	10.2%	9.1%	3.66%
no. species not listed	0	3	0	13	0	0	0	0	0	434	86	9	0	0	4	3	0	0	5	8	0	0					

the number of peptide variants for the spectra that were identified in a particular sample.

The support for a particular taxa, e.g., taxa A, was calculated as:

(Equation 1)

$$\%\ spectral\ count\ matching = \left(\frac{sum\ of\ spectra\ matching\ to\ taxon\ A}{total\ spectral\ count\ of\ sample}\right) \times 100$$

This statistic includes the relative abundance of each peptide spectral match and is used to compare the likelihood that a given peptide matches to 'taxon A'.

Results

Hemoglobin and albumin were detected in all body parts immediately after a blood meal (0 wk A and 0 wk B). In addition, the high number of spectral counts matching the expected mouse blood source at 0 weeks confirms the efficacy of LC-MS/MS identification. By 8 weeks post-blood meal, we were still able to detect hemoglobin and albumin in several insect parts, but in

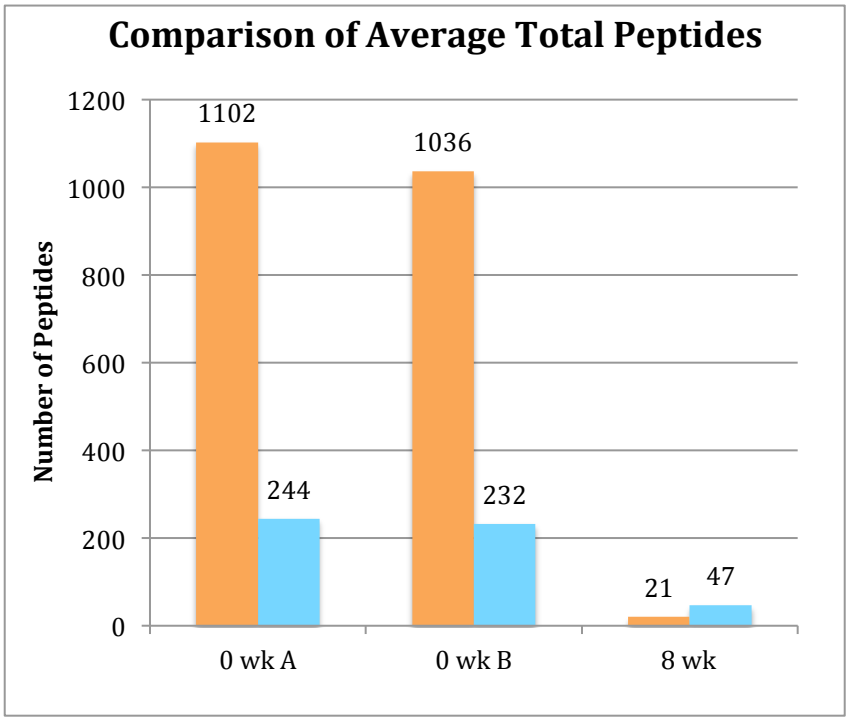
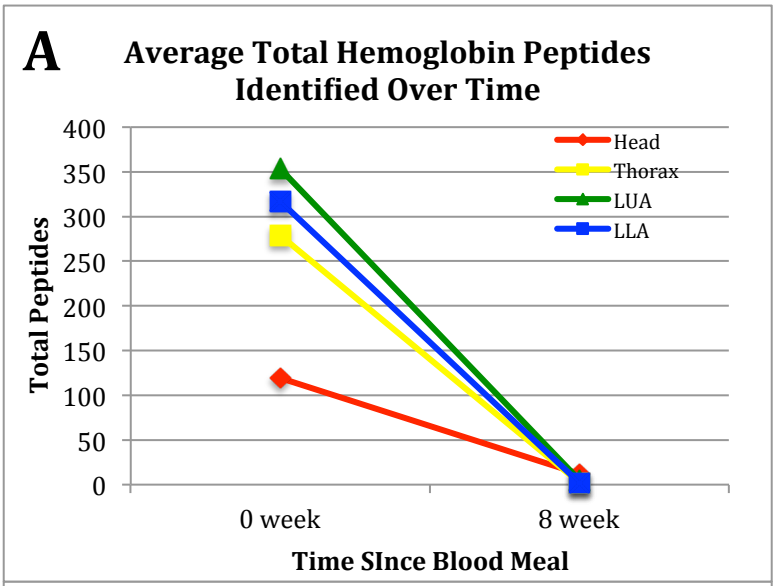


Figure 4. Plot comparing the total number of peptides detected for each of 3 insects over 2 time points. Orange= hemoglobin, Blue= albumin.

lower spectral count quantities. The details supporting these conclusions are described below.

Detection of mouse hemoglobin and albumin peptides

Analysis of 0wk A and 0wk B demonstrate detection of similar numbers of both hemoglobin and albumin peptides in each body part (**Appendix B).** Approximately 4.5x more hemoglobin peptides than albumin peptides were detected in these bugs (**Figure 4**). The consistency of spectral totals between both 0 wk bugs exists to the level of bug parts as well. For both 0wk bugs A and B, the highest spectral counts of hemoglobin were found in the upper abdomen (415 peptides and 292 peptides respectively) and the lowest levels in the head (50, 189 peptides respectively) (**Figure 5**; **Appendix B**).



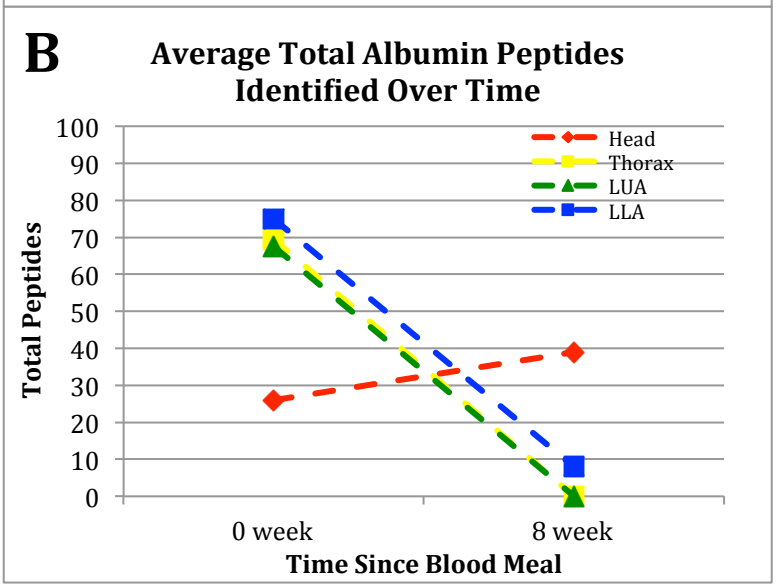


Figure 5. Preliminary time course comparison of the total peptides of (A) hemoglobin and (B) albumin present in insect body parts over time. 0 wk insects (N=2) were averaged to provide the total peptide count of both Hb and Alb. The four lines represent change in the number of peptides detected in each body part at two time points.

Even after 8 weeks, low levels (2-12 peptide spectral counts) of hemoglobin peptides were detected in all insect body parts, however at 8 wk the highest spectral counts were from the head (**Figure 6**). Albumin was only found in two body parts of the 8wk bug, the head (spectral count 39) and lower abdomen (spectral count 8).

Peptide Identification

Although we searched the peptide sequences for known hemoglobin and albumin peptides in *any* vertebrate taxa, our results matched the expected of *M. musculus* with high fidelity. Our stringent spectra filters resulted in no reverse database matches and thus gave peptide false discovery rates of 0 with the power to detect less than 0.01%. **Table 1** shows the pipeline summary for sample 0 wk B-head. Note that over 95% of the 191 spectra from this sample matched to peptides previously reported in *M. musculus.* The percent of spectral counts matching are higher than the percent peptides identified matching because the former includes the relative abundance of each peptide spectral match, and the peptide variants that do not match tend to be of low abundance.

Although for both 0 wk samples some of the non-redundant peptides identified did not match *M. musculus,* in all cases the percent spectral counts matching mouse were greater than 82% (**Appendix B**). Hemoglobin peptides of 0wk insect parts matched correctly to *M. musculus* between 82% and 95.2%. Peptide variants accounted for the reduction from 100%, yet all variants are attributable to repeated versions of the same six peptides (**Table 2**).

The 8 wk bug parts matched correctly to *M. musculus* between 0% and 100%, but only 8 unique hemoglobin peptides and 4 unique albumin peptides were identified.

Table 2. List of all *M. musculus* alpha and beta hemoglobin peptides identified by LC-MS/MS spectral matching. Peptide variant amino acid sequences are not listed, but the number of unique peptide variants that align to *M. musculus* is listed, and peptides are sorted by highest number of unique non-matching peptide variants

Hemoglobin Peptide	# Unique Non-Matching Peptide Variants
IGGHGAEYGAEALER	5
VNSDEVGGEALGR	4
VADALASAAGHLDDLPGALSALSDLHAHK	4
VVAGVATALAHK	1
TYFPHFDVSHGSAQVK	1
LLGNMIVIVLGHHLGK	1
VLSGEDKSNIK	0
MFASFPTTK	0
LRVDPVNFK	0
LLSHCLLVTLASHHPADFTPAVHASLDK	0
FLASVSTVLTSKYR	0
VHLTDAEK	0
AAVSCLWGK	0
LLVVYPWTQR	0
YFDSFGDLSSASAIMGNAK	0
VITAFNDGLNHLDSLK	0
GTFASLSELHCDK	0
LHVDPENFR	0
DFTPAAQAAFQK	0

Discussion

Using LC-MS/MS, I detected and identified host blood proteins in all four insect body parts, left lower abdomen (LLA), left upper abdomen (LUA), thorax and head, immediately after the blood meal was taken. In a single insect, up to 1102 spectra matched hemoglobin peptides and 244 matched albumin peptides. Although the total spectral count detectable decreased to less than 50 total peptides by 8 weeks, there were still blood proteins present in the head and LLA (**Appendix B**). Hemoglobin and albumin were detected in all body parts immediately after a blood meal (0 week). These results combined with the preliminary time course analysis

suggests promise in utilizing this methodology for experimental insects from the field.

Percent spectral count matching

Accurate identification of a blood meal source is critical to ensuring the usefulness of this methodology for field-collected insect vectors. In all freshly fed body part samples, the majority of the peptide variant matched *M. musculus* hemoglobin. Since we constructed a database including hemoglobin proteins from all taxa in GenBank, we are able to identify the taxonomy of the non-matches and are likely to find non-matches in field-collected insects. In total, over the three insects and four body parts, 2 - 20 % of the peptides identified had not previously been reported in mouse.

In some cases, variant peptides not matching *M. musculus* differ at only one amino acid, and matched to other *Mus* species, suggesting potential polymorphisms not documented in Genbank (Benson et al., 2007). The same peptide variants were present in multiple body parts of multiple samples, but never lowered the percent hemoglobin spectral count matching below 82.00% in freshly fed insects and 60% in the 8 wk insect (**Appendix B**). We did not further investigate the reason that these few peptides do not match, because in spite of these non-matches we demonstrate the ability to accurately identify blood meal sources in future field experiments.

Insect body part analysis

The variation among body parts in the number of spectral counts matching hemoglobin or albumin shows that the common practice of using the same body part for comparative blood meal analysis is important. DNA analyses routinely use the abdomen (e.g., Stevens et al., 2012; Dorn et al., 2014, Pizarro and Stevens 2009), while some antibody based analyses use the thorax (Gürtler et al., 2014). This appears to be the first study comparing the ability to detect blood meals from multiple body parts. In particular, little if any research has been done on the ability to detect blood meals in the insect vector head. Compared to the other body parts, few peptides were detected in the head of freshly fed insects, but more peptides were detected in the head at 8 weeks than any other body part (**Figure 5**).

Spectral counts of hemoglobin in freshly fed insects followed the trend proposed by hypothesis 1, that peptides would be found in all consecutive body parts, with fewest in the head. The thorax, LUA and LLA had up to 8 times more spectral matches than the head (**Appendix B**). The insect abdomen is known to contain a large volume of blood immediately after a blood meal, which is represented by the high levels of hemoglobin peptides found there (Lehane et al., 1991). 0 wk insect abdomen had fewer albumin peptides detected than hemoglobin, supporting hypothesis 2 that hemoglobin would be more highly detectable than albumin. The thorax, LUA and LLA had 2-5 times more albumin peptides detected than the head, further demonstrating the abundance of peptides detectable in the lower digestive tract immediately after a blood meal (**Appendix B**). It is important to note that because each abdomen section studied represents about one-fourth of

the insect abdomen, about 1000 to 1400 peptides may have been detected if the entire abdomen was used (**Appendix B**; **Figure 4**).

Results of the peptide spectral count in the thorax, LUA and LLA at 8 weeks support hypothesis 3 that fewer peptides will be detectable over time. Hemoglobin peptides were detected in these 3 insect parts but the spectral count decreased to 5 peptides in the LUA and 2 peptides in the thorax and LLA, a greater than 100 fold decrease for both parts. Detectability of albumin in these parts also decreased after 8 weeks, to 8 peptides in the LLA and 0 in the thorax and LUA. These data support a similar trend in the digestion and resultant decrease in detectability of both hemoglobin and albumin in the thorax and abdomen over 8 weeks.

If the results of the peptide spectral count in the head at 8 weeks are repeatable, it would indicate that blood meals may be broken down at different rates in different parts of the body. At 8 weeks after taking a blood meal, quite a few more hemoglobin and albumin peptides were detected in the insect head (12 and 39 peptides respectively) than other body parts. Although the other body parts support hypothesis 2 that hemoglobin is more stable than albumin, and hypothesis 1 that freshly fed vectors will have more protein detectable, further study with additional bugs is needed to validate this surprising result (**Appendix B**). If this observation holds through additional sampling, possible explanations are that head results may be skewed due to the low quantity of insect peptides from the relatively small size of the head or that blood meal remnants in the crop or esophagus within the foregut are not subject to digestive enzymes found in the gut, increasing the relative spectral count in the insect head over time (Lehane et al., 1991).

Despite surprising results of the detectability of blood meal peptides in the head at 8 weeks, consistency of spectral counts in the freshly fed insect thorax and abdomen supports the idea of using these body parts for time course analysis.

Expansion of time course analysis

These data support hypothesis 1 that hemoglobin from a vertebrate blood meal passes through the insect and traces of identifiable protein remain in the lower digestive tract. These data also support hypothesis 2 that hemoglobin is more stable than albumin and identifiable in the insect gut at higher levels than albumin. Hypothesis 3 that there is a time threshold up to which the vertebrate blood meal source can be identified, was somewhat supported as the overall levels of detectable peptides decreased over time; however even at 8 weeks peptides of hemoglobin were detected in all parts and albumin was detected in the head and lower abdomen.

The large difference in detectability between hemoglobin and albumin coupled with the spectral count consistency of the replicate freshly fed insects indicate this is a promising approach and more time points with more replicates per time point should be investigated. With sufficient time points and replicates per time, one could develop a multivariate interpolation method for blood meal analysis using two proteins, hemoglobin and albumin., multiple sampling times, from 0 to over 8 weeks, and initially 3-5 replicates per sampling time. The multivariate analysis would provide a reference for estimating the time since the last blood meal.

Based on this study, the abdomen yields the highest spectral count of both proteins at 0 wk and both LUA and LLA follow similar trends over the two time points. In theory, statistics from a multivariate analysis of controlled feedings and time sampling would allow one to estimate the time since last feeding, and the error in that estimate. Knowing when insects feed is useful for vector control because we can identify when and how often vectors feed as well as potential movement into and out of houses. Beyond attempting to isolate the blood meal source to determine the most fitting ecohealth approach, knowing when to implement vector control may save costs of repetitive insect control like spraying pesticide (Monroy et al., 2009). Furthermore, comparing relative peptide quantities of more than one matching blood source may allow us to place the estimated feeding times of multiple blood meals on a timeline (Önder et al., 2013). However, repeated feeding on the same blood meal source and perhaps the size of each blood meal may complicate such analyses.

The consistency of spectral counts at measured time points after blood meals are likely to vary based on many factors including quantity of the blood meal and difference in digestion. Even with variation due to uncontrollable factors, analyzing insects from the field may allow identification of the time since a blood meal and perhaps even the sequence of multiple blood meals based on relative quantities of identified peptides and multivariate interpolation.

Future Directions

In a future study, Triatomine insects could be collected from the field in regions that are at risk of Chagas disease and where the disease is currently classified as endemic. These insects could be dissected similarly to the schematic shown in **Figure 2** and analyzed for the presence of vertebrate blood proteins, hemoglobin and albumin. If the insect had taken a blood meal within recent weeks and potentially recent months, LC-MS/MS based peptide identification could show the "highest likelihood host match" from percentage spectral count matching, as done in this study. The three insects in this study only fed on *M. musculus* as the vertebrate host and all freshly fed insects were best matched with *M. musculus* by the calculation shown in Equation 1. At 8 weeks, this study suggests that hemoglobin in all body parts is a useful indicator of vertebrate host, but the absence of albumin peptides in the thorax and LUA suggest more difficulty in detecting enough albumin peptides at 8 weeks to identify the host. However, when examining an insect from the field, even a family or genus level identification may be sufficient in guiding an ecohealth disease prevention measure, and may be used to identify human risk of disease.

When the collection locality of the insect is known, it is likely that a limited number of species from a given vertebrate genus are possible hosts. For example, of the seven species within the genus *Felis*, only the house cat, *F. catus* is commonly found in the Americas [retrieved April 18, 2017, from the Integrated Taxonomic Information System on-line database, http://www.itis.gov]. Identifying *Felis* blood proteins in Triatomine insects would suggest that domestic pets are the vertebrate

blood meal host. Targeting ecohealth methods to include the education of pet owners and preventing contact of house cats with potential vector hiding places outside the home might be a precise and beneficial approach. Previous research by Monroy et al. has shown that recognizing the potential diversity of blood meal sources is critical in designing and implementing ecohealth strategies (Monroy et al., 2009)

In conclusion, when considering the potential benefits from ecohealth approaches, the need for elucidating the Chagas disease transmission pathway becomes evident. Based on the success of LC-MS/MS detection and identification in this study, the future looks bright for creating and utilizing a control time course analysis of hemoglobin and albumin. With this, we may continue to fend off the spread of infection and alleviate the global burden of Chagas disease.

Acknowledgments

This work was supported by NIH Grant 8P20GM103449 from the INBRE program of the NIGMS and by NSF grant BCS-1216193 as part of the joint NSF-NIH-USDA Ecology and Evolution of Infectious Diseases program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Institutes of Health or the National Science Foundation.

I would like to thank Justin Schmidt, Southwestern Biological Institute in Tucson, AZ for providing the insects; the UVM Proteomics Core and particularly Bethany Ahler for assisting with sample processing; Dr. Stevens' Lab at UVM for guiding my research; and Judith Keller for all the help along the way.

Appendix

Appendix A. Pivot table for 0 wk B hemoglobin head. This is an example of the taxonomic data management used for each body part (head thorax, left upper abdomen, left lower abdomen) for both hemoglobin and albumin for each of 3 insects. Row labels in the far left column are collapsible and all fall within one peptide identified for the sample. Distinct counts of taxonomic levels that match to the peptide using our database are represented by the other columns.

Row Labels	Distinct Count of class	Distinct Count of order	Distinct Count of family	Distinct Count of genus	Distinct Count of species
DFTPAAQAAFQK	1	2	4	5	13
Mammalia	1	2	4	5	13
Insectivora	1	1	1	1	2
Erinaceidae	1	1	1	1	2
Erinaceus	1	1	1	1	2
Erinaceus europaeus	1	1	1	1	1
western European hedgehog	1	1	1	1	1
Erinaceus sp.	1	1	1	1	1
	1	1	1	1	1
Rodentia	1	1	3	4	11
Cricetidae	1	1	1	1	2
Microtus	1	1	1	1	2
Microtus ochrogaster	1	1	1	1	1
prairie vole	1	1	1	1	1
Microtus pennsylvanicus	1	1	1	1	1
meadow vole	1	1	1	1	1
Muridae	1	1	1	1	6
Mus	1	1	1	1	6
Mus cookii	1	1	1	1	1
Cook's mouse	1	1	1	1	1
Mus macedonicus	1	1	1	1	1
Macedonian mouse	1	1	1	1	1
Mus minutoides	1	1	1	1	1
Southern African pygmy mouse	1	1	1	1	1
Mus musculus	1	1	1	1	1
house mouse	1	1	1	1	1

Appendix B. Summary of results table. LC-MS/MS spectral counts of hemoglobin (Hb) and albumin (Alb) peptides found in each body part of 3 experimental insects. 0 wk A and B are freshly fed bugs and 8 wk is a bug sampled 8 weeks post-blood meal. Body parts LUA and LLA refer to the left upper abdomen and left lower abdomen, respectively. Percent spectral count matching is a weighted measure referring to either the number of Hb or Alb peptides that matched to mouse divided by the total spectral count for the sample.

Protein	Insect	Body Part	Spectral Count	Percent spectral count matching *Mus musculus*
Hemoglobin	7	Head	50	82.00%
		Thorax	272	86.76%
		LUA	415	85.06%
		LLA	365	83.29%
	72	Head	189	95.24%
		Thorax	286	86.71%
		LUA	292	84.93%
		LLA	269	86.62%
	48	Head	12	100.00%
		Thorax	2	100.00%
		LUA	5	60.00%
		LLA	2	100.00%
Albumin	7	Head	18	94.44%
		Thorax	75	93.33%
		LUA	57	98.25%
		LLA	94	90.43%
	72	Head	34	88.24%
		Thorax	64	96.88%
		LUA	78	97.44%
		LLA	56	98.21%
	48	Head	39	100.00%
		Thorax	0	-
		LUA	0	-
		LLA	8	0.00%

Works Cited

Ballif, Bryan A., Zhongwei Cao, Daniel Schwartz, Kermit L. Carraway, and Steven P. Gygi. "Identification of 14-3-3Îµ Substrates from Embryonic Murine Brain." *Journal of Proteome Research* 5.9 (2006): 2372-379. Web.

Ballif, Bryan A, GR Carey, S Sunyaev, SP Gygi, Large-Scale Identification and Evolution Indexing of Tyrosine Phosphorylation Sites from Murine Brain. *Journal of Proteome Research*, 2008, 7:311-8. Web.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. Nucleic Acids Res. 2009 Jan;37(Database issue):D26-31. doi: 10.1093/nar/gkn723. Epub 2008 Oct 21. PMID: 18940867

Boscardin, Silvia B, Ana Torrecilhas Claudia Troccoli, Romina Manarin, Silvia Revelli, Elena Rey Gonzalez, Renata Tonelli Rosito, and Ariel Silber Mariano. "Chagas' Disease: An Update on Immune Mechanisms and Therapeutic Strategies." *Journal of Cellular and Molecular Medicine* 14.6b (2010): 1373-384. Web.

Brasil, Pedro Eaa, Liane De Castro, Alejandro M. Hasslocher-Moreno, Luiz Hc Sangenis, and José U. Braga. "ELISA versus PCR for Diagnosis of Chronic Chagas Disease: Systematic Review and Meta-analysis." *BMC Infectious Diseases* 10.1 (2010): n. pag. Web.

Coronado, X., M. Rozas, C. Botto-Mahan, S. Ortiz, P. E. Cattan, and A. Solari. "Molecular Epidemiology of Chagas Disease in the Wild Transmission Cycle: The Evaluation in the Sylvatic Vector Mepraia Spinolai from an Endemic Area of Chile." *American Journal of Tropical Medicine and Hygiene* 81.4 (2009): 656-59. Web.

Goloborodko AA, Levitsky LI, Ivanov MV, Gorshkov MV. J Am Soc Mass Spectrom. 2013 Feb;24(2):301-4. doi: 10.1007/s13361-012-0516-6. Epub 2013 Jan 5. PMID: 23292976

Gullan, P. J., P. S. Cranston, and Karina H. McInnes. *The Insects: An Outline of Entomology*. Chichester, West Sussex: Wiley Blackwell, 2014. Print.

Gürtler, Ricardo E., María C. Cecere, Gonzalo M. Vázquez-Prokopec, Leonardo A. Ceballos, Juan M. Gurevitz, and María Del Pilar Fernández. "Domestic Animal Hosts Strongly Influence Human-Feeding Rates of the Chagas Disease Vector Triatoma Infestans in Argentina." *PLoS Neglected Tropical Diseases* 8.5 (2014): n. pag. Web.

Howard, Elizabeth J., Xu Xiong, Yves Carlier, Sergio Sosa-Estani, and Pierre Buekens. "Frequency of the Congenital Transmission of Trypanosoma Cruzi: A Systematic Review and Meta-Analysis." *British Journal of Obstetrics and Gynocology* 121.1 (2013): 22-23. Web.

Keller, J. I, C. Monroy, J. Vincent, R. M. St. Clair, L. Stevens, B. A. Ballif. 2017. Chagas Disease Vector Blood Meal Sources Identified by Protein Mass Spectrometry. PLoS Journal of Neglected Tropical Disease **[in prep]**

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. Am J Med. 2014;127(5):421-6.

Laskay ÜA, Breci L, Vilcins IM, Dietrich G, Barbour AG, Piesman AG, Wysocki VH. Survival of host blood proteins in Ixodes scapularis (Acari: Ixodidae) ticks: a time course study. 2013; J Med Entomol 50(6):1282-90.

Laskay ÜA, Burg J, Kaleta EJ, Vilcins IE, 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; Biol Chem (393):195-201.

Lee, Bruce Y., Kristina Bacon M., Maria Bottazzi Elena, and Peter Hotez J. "Global Economic Burden of Chagas Disease: A Computational Simulation Model." *The Lancet Infectious Diseases* 13.4 (2013): 342-48. Web.

Lehane, Mike J. *Biology of Blood-Sucking Insects.* London: Harper Collins Academic, 1991. Print.

Life cycle schematic of protozoan parasite that can be transmitted to humans by blood-sucking triatomine bugs Trypanosoma cruzi. Digital image. *CDC.gov*. Center for Disease Control Division of Parasitic Diseases and Malaria, 3 May 2016. Web.

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.

Manarin, Romina, María Lamas C., Emanuel Bottasso, Esteban Serra, Silvia Revelli, and Claudio Salomón J. "Efficacy of Novel Benznidazole Solutions during the Experimental Infection with Trypanosoma Cruzi." *Parasitology International* 62.1 (2013): 79-81. Web.

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

Monroy, Carlota, Dulce Maria Bustamante, Sandy Pineda, Antonieta Rodas, Xochitl Castro, Virgilio Ayala, Javier Quiñónes, and Bárbara Moguel. "House Improvements and Community Participation in the Control of Triatoma Dimidiata Re-infestation in Jutiapa, Guatemala." *Cadernos De Saúde Pública* 25 (2009): n. pag. Web.

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.

Oshaghi MA, Chavshin AR, Vatandoost H, Yaaghoobi F, Mohtarami F, Noorjah N. Effects of post-ingestion on physical conditions on PCR amplification of host blood meal DNA in mosquitoes. 2006; Exp Parasitol 112(4):232-6.

Ozaki, Yaeko, Maria Elena Guariento, and Eros Antonio De Almeida. "Quality of Life and Depressive Symptoms in Chagas Disease Patients." *Quality of Life Research* 20.1 (2011): 133-38. Web.

Pizarro, Juan Carlos, and Lori Stevens. "A New Method for Forensic DNA Analysis of the Blood Meal in Chagas Disease Vectors Demonstrated Using Triatoma Infestans from Chuquisaca, Bolivia." *PLoS ONE* 3.10 (2008): n. pag. Web.

Salomon, Claudio J. "First Century of Chagas' Disease: An Overview on Novel Approaches to Nifurtimox and Benznidazole Delivery Systems." *Journal of Pharmaceutical Sciences J. Pharm. Sci.* 101.3 (2011): 888-94. 12 Dec. 2011. Web.

Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V, Mizrachi I, Ostell J, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Yaschenko E, Ye J.Nucleic Acids Res. 2009 Jan;37(Database issue):D5-15. doi: 10.1093/nar/gkn741. Epub 2008 Oct 21. Erratum in: Nucleic Acids Res. 2009 May;37(9):3124. PMID: 18940862

Sosa-Estani, Sergio, Lisandro Colantonio, and Elsa Leonor Segura. "Therapy of Chagas Disease: Implications for Levels of Prevention." *Journal of Tropical Medicine* 2012 (2012): 1-10. Web.

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 Infect Dis. 2012 Apr;18(4):646-9.

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.

Waleckx, Etienne, Julianne Suarez, Bethany Richards, and Patricia L. Dorn. "Triatoma SanguisugaBlood Meals and Potential for Chagas Disease, Louisiana, USA." *Emerging Infectious Diseases* 20.12 (2014): 2141-143.

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).