



OPEN ACCESS

Citation: Keller JI, Ballif BA, St. Clair RM, Vincent JJ, Monroy MC, Stevens L (2017) Chagas disease vector blood meal sources identified by protein mass spectrometry. PLoS ONE 12(12): e0189647. https://doi.org/10.1371/journal.pone.0189647

Editor: Sebastien Gourbiere, Universite de

Perpignan, FRANCE

Received: September 7, 2017

Accepted: November 29, 2017

Published: December 12, 2017

Copyright: © 2017 Keller et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information

files.

Funding: This work was supported by National Institute of Health (http://www.nih.gov/) grant 8P20GM103449 (BAB) from the INBRE program of the NIGMS, National Institutes of Health (NIH) (http://www.nih.gov/) grant R03Al26268/1-2 (LS) from the National Institute of Allergy and Infectious Diseases (NIAID) (https://www.niaid.nih.gov/) and National Science Foundation (http://nsf.gov/) grant BCS-1216193 (LS, CM) as part of the joint NSF-

RESEARCH ARTICLE

Chagas disease vector blood meal sources identified by protein mass spectrometry

Judith I. Keller¹, Bryan A. Ballif¹*, Riley M. St. Clair¹, James J. Vincent¹, M. Carlota Monroy^{1,2}, Lori Stevens¹*

- 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, Ciudad de Guatemala, Guatemala
- * lori.stevens@uvm.edu (LS); bballif@uvm.edu (BAB)

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.



NIH-USDA Ecology and Evolution of Infectious Diseases program. 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. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

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.

Methods

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

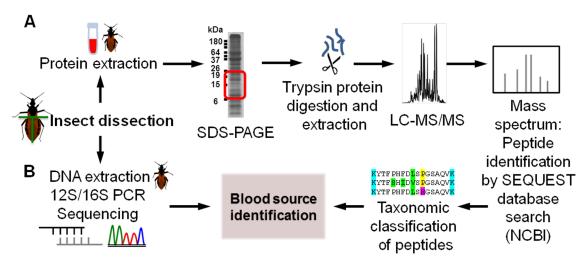


Fig 1. General workflow describing LC-MS/MS-based and DNA-based identification of blood meal sources. Insects were dissected into left and right abdomen, and hemoglobin peptides from (A) mouse blood as well as triatomine insect vectors were identified with LC-MS/MS, while (B) triatomine insect vector blood meals were also identified with DNA-based methods for comparison.



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.

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 K₂EDTA; 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.

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.

Protein extraction, SDS-PAGE, and mass spectrometry

We extracted proteins from $\it M. musculus$ 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

Table 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		
			Village	Subvillage	Ecotope	
FER 0051-1	LC-MS/MS	М	Azacualpa	El Tablón	domestic	alive
FER 0051-2	DNA					
FER 0076-1	DNA	М	M Azacualpa	Ocotillo	domestic	dead
FER 0076–2	LC-MS/MS					
FER 0101-1	DNA	М	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					

^{1,} left abdomen; 2, right abdomen; M, male; F, female.



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 μ l of mouse blood was run and analyzed per gel lane. For insect samples, proteins were extracted by adding 100 μ l 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 $20~\mu$ l 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.

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)

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 Δ Cn 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 (Fig 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: 1996 non-redundant peptides identified in sample																	
A M. musculus NP_032244.2, BAG16710.1 no. amino acids/peptide	ந alpha_17-31	o alpha_32-40	ട്ട alpha_41-56	o alpha_91-99	ਨੇ alpha_128-139	ο beta_1-8	ದ beta_18-30	16	16 netg	1 beta_66-82	ධ beta_83-95	16	100-120 100-120	പ്പ് beta_121-132	ದ beta_133-144	다 beta_133-146	Total 214
no. peptide variants	1	1	1	1	1	1	1	2	2	1	1	2	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	Х	Х	X	Х	Х	Х	Х	Х	X	Х	Х	Χ	Х	Х	Х	Χ	
Mus spicilegus							Х	Х	X	X	Х	X	X	Х	Х	X	
Mus macedonicus							X	Х	X		X	X	Х	X	X	X	
Mus spretus								X		X	X	X		X			
Cricetomys gambianus	X	X	X	X	X												
Mus pahari										X					Х	X	
no. species not listed	5		15	434	88	9		4	4		28	4		7	38	37	

Fig 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.

https://doi.org/10.1371/journal.pone.0189647.g002

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 <u>Fig 2</u> and <u>Table 2</u> for more detail).

Table 2. Using the information from the pipeline (Fig 2), two summary statistics quantify the strength of blood meal source identification.

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%

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).



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].

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.

Mouse blood

We were able to identify hemoglobin protein peptides unambiguously with our proteomics-based, LC-MS/MS approach (Fig 3, Table 3). 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 variants 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)



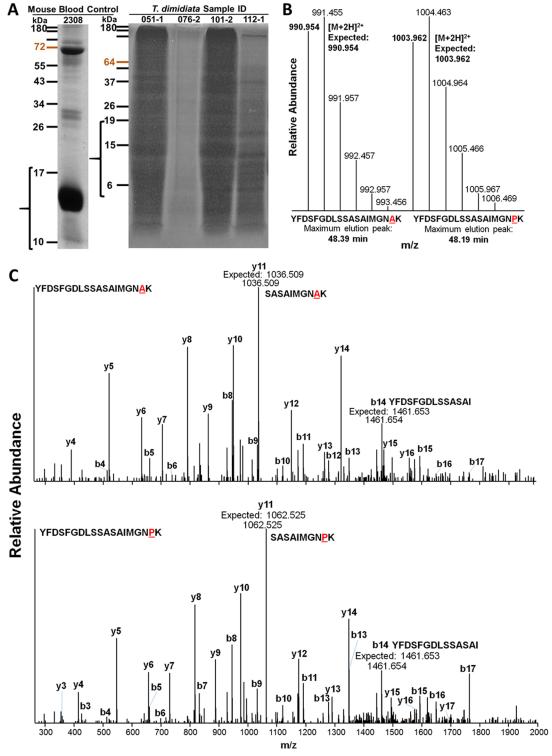


Fig 3. Proteomics-based, LC-MS/MS distinguishes unique but nearly identical peptides. (A) Representative SDS-PAGE results of mouse blood samples and of four *T. dimidiata* insect vectors. (B) MS1 spectra of two doubly-charged peptide ions differing by a single amino acid-YFDSFGDLSSASAIMGNAK and YFDSFGDLSSASAIMGNPK, 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 S1 and S2 Tables.

https://doi.org/10.1371/journal.pone.0189647.g003

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).

Table 3. Tryptic peptides identified in mouse blood control samples, including hemoglobin position and spectral counts.

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		

a indicates peptides not matching the known blood sources, M. musculus; 1996, 100 mM EDTA; 2308, 18 mM K₂EDTA; 2310, no preservative



Table 4. Summary statistics and blood source identifications of mouse blood controls and triatomine insect samples.

Sample	Blood source identification	Percent peptides identified matching	Percent spectral count matching	Percent amino acids identified matching	
Mouse b	lood 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%	
101	C. lupus	92.3%	95.7%	99.5%	
112	C. lupus	100.0%	100.0%	100%	

1996, 100 mM EDTA; 2308, 18 mM K₂EDTA; 2310, no preservative

https://doi.org/10.1371/journal.pone.0189647.t004

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).

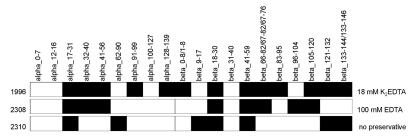


Fig 4. Hemoglobin protein peptide coverage from identified tryptic peptides for three mouse blood samples. Three mouse blood samples stored in varying amounts of EDTA are shown. The peptides are shown with equal width, not in proportion to the length of the peptide, and only peptides greater than four amino acids are shown.



Table 5. Peptides identified in triatomine insects, including hemoglobin position and spectral counts.

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			

^a indicates peptides not matching the most likely blood meal (051, cat; 076, 101, 112, dog).

https://doi.org/10.1371/journal.pone.0189647.t005

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

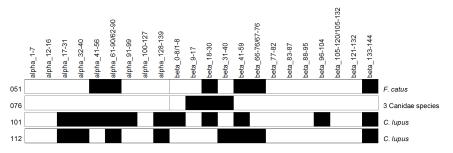


Fig 5. Hemoglobin protein peptide coverage from identified tryptic peptides from blood meals of four triatomine insects. Triatomine insect vectors were collected alive (051, 101, 112) and dead (076). The peptides are shown with equal width, not in proportion to the length of the peptide, and only peptides greater than four amino acids are shown.

^b indicates a peptide that was added to GenBank 29 Dec 2016, after the original database was created.



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.

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).

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

Table 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



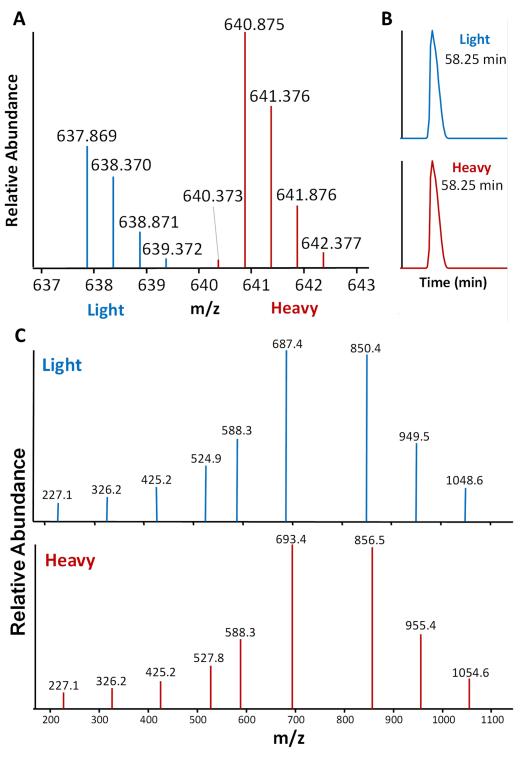


Fig 6. Stable isotope-labeled standard facilitates quantification of native hemoglobin peptide. (A) MS1 spectrum showing the relative abundance of a native hemoglobin peptide derived from the blood meal of a field-collected kissing bug (blue isotopic envelope) in comparison to an introduced synthetic 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.



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.

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 (IGGHGAEY-GAEALER) 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%).

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, VADALTTAVSHIDDLPGALSALS DLHAYK 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 Gen-Bank 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 multiuser 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 (\$3 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.

Supporting information

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

(PDF)

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

(PDF)

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

(PDF)

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

(PDF)

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

(PDF)



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

(PDF)

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

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. (PDF)

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. (PDF)

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). (PDF)

Acknowledgments

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.

Author Contributions

Conceptualization: Judith I. Keller, Bryan A. Ballif, M. Carlota Monroy, Lori Stevens.

Data curation: Judith I. Keller, M. Carlota Monroy, Lori Stevens.

Formal analysis: Judith I. Keller, Bryan A. Ballif, Lori Stevens.

Funding acquisition: Bryan A. Ballif, M. Carlota Monroy, Lori Stevens.

Investigation: Judith I. Keller, Bryan A. Ballif, Riley M. St. Clair, Lori Stevens.

Methodology: Judith I. Keller, Bryan A. Ballif, Lori Stevens.

Project administration: Judith I. Keller, Bryan A. Ballif, M. Carlota Monroy, Lori Stevens.

Resources: Bryan A. Ballif, M. Carlota Monroy, Lori Stevens.

Software: Judith I. Keller, Bryan A. Ballif, James J. Vincent, Lori Stevens.

Supervision: Bryan A. Ballif, M. Carlota Monroy, Lori Stevens.



Validation: Judith I. Keller, Bryan A. Ballif, Lori Stevens.

Visualization: Judith I. Keller, Bryan A. Ballif, Lori Stevens.

Writing – original draft: Judith I. Keller, Lori Stevens.

Writing - review & editing: Judith I. Keller, Bryan A. Ballif, M. Carlota Monroy, Lori Stevens.

References

- 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): https://doi.org/10.1098/rstb. 2013.0552 PMID: 25688013
- Cohen JE, Gürtler RE. Modeling household transmission of American trypanosomiasis. Science (New York, NY). 2001; 293(5530):694–8. https://doi.org/10.1126/science.1060638 PMID: 11474111
- 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.
- 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. https://doi.org/10.1128/CMR.00005-11 PMID: 21976603
- 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. https://doi.org/10. 1371/journal.pntd.0005422 PMID: 28426735
- 6. Rassi A, Marin-Neto JA. Chagas disease. The Lancet. 2010; 375(9723):1388–402.
- 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. https://doi.org/10.1371/journal.pntd.0002300 PMID: 24205411
- Sánchez-Valdéz FJ, Pérez Brandán C, Ferreira A, Basombrío MÁ. Gene-deleted live-attenuated *Trypa-nosoma cruzi* parasites as vaccines to protect against Chagas disease. Expert Review of Vaccines. 2015; 14(5):681–97. https://doi.org/10.1586/14760584.2015.989989 PMID: 25496192
- Coura JR, De Castro SL. A critical review on Chagas disease chemotherapy. Memórias do Instituto Oswaldo Cruz. 2002; 97(1):3–24. PMID: 11992141
- Justi SA, Galvão C. The Evolutionary Origin of Diversity in Chagas Disease Vectors. Trends in Parasitology. 2017; 33(1):42–52. https://doi.org/10.1016/j.pt.2016.11.002. PMID: 27986547
- 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. https://doi.org/10.1371/journal.pntd.0003365 PMID: 25474154.
- 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. https://doi.org/10.3201/eid2012.131576 PMID: 25418456
- 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. https://doi.org/10.4269/ajtmh.12-0458 PMID: 23382165
- 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. https://doi.org/10.4269/ajtmh.12-0448. PMID: 23382173
- 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. PMID: 19287859
- 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. https://doi.org/10.1089/vbz.2007.0106 PMID: 18021027
- 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. https://doi.org/10.2987/09-0012.1 PMID: 19653507



- 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. 1475-9292-6-1 https://doi.org/10.1186/1475-9292-6-1 PMID: 17222347
- 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. https://doi.org/10.1371/journal.pone.0003585 PMID: 18974787
- 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. https://doi.org/10.1371/journal.pntd. 0003047 PMID: 25211347
- 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. https://doi.org/10.1016/j.amjmed.2013.12.017 PMID: 24398362
- 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. https://doi.org/10.3201/eid1804.111396 PMID: 22469536
- Peña VH, Fernández GJ, Gómez-Palacio AM, Mejía-Jaramillo AM, Cantillo O, Triana-Chávez O. Highresolution 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. https://doi.org/10.1371/journal.pntd.0001530 PMID: 22389739
- 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. https://doi.org/10.1016/j.meegid.2013.07.002 PMID: 23856437
- 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. PMID: 16474087
- 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. PMID: 17915521
- 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. https://doi.org/10.1111/j. 1755-0998.2008.02469.x PMID: 21564560
- 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.
- 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. https://doi.org/10.1038/ncomms2730 PMID: 23612287
- 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. PMID: 12693851
- Souza Rd CMd, 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. PMID: 21340350
- 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. PMID: 11296839
- 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. https://doi.org/10.4269/ajtmh.2010.09-0397 PMID: 20348502
- **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. PMID: 15605639



- 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. PMID: 15654457
- 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. PMID: 10973153
- Vasquez AMd, Samudio FE, Saldaña A, Paz HM, Calzada JE. Eco-epidemiological aspects of *Trypano-soma cruzi*, *Trypanosoma rangeli* and their vector (*Rhodnius pallescens*) in Panama. Revista do Instituto de Medicina Tropical de São Paulo. 2004; 46(4):217–22. PMID: 15361974
- 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. https://doi.org/10.1073/pnas.1310885110 PMID: 24127577
- Laskay ÜA, Burg J, Kaleta EJ, Vilcins I-ME, Telford SR Iii, Barbour AG, et al. Development of a host blood meal database: de novo sequencing of hemoglobin from nine small mammals using mass spectrometry. 2012.
- 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. PMID: 24843933
- 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. https://doi.org/10.1007/s13361-015-1185-z PMID: 26111519
- 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. https://doi.org/10.1093/nar/gkn723 PMID: 18940867.
- 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. https://doi.org/10.1083/jcb.200802105 PMID: 18591425
- 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. https://doi.org/10.1038/nrendo.2011.234 PMID: 22290358
- 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. https://doi.org/10.1021/pr0701254 PMID: 18034455
- 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. https://doi.org/10.1093/molbev/msu403
 PMID: 25556236
- **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. https://doi.org/10.1073/pnas.0832254100 PMID: 12771378
- 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. https://doi.org/10.1002/pmic.201200003 PMID: 22807455
- Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. Nature Reviews Molecular Cell Biology. 2004; 5(9):699–711. https://doi.org/10.1038/nrm1468 PMID: 15340378
- 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. https://doi.org/10.1007/s13361-012-0516-6 PMID: 23292976
- 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. https://doi.org/10.1093/nar/gkn741 PMID: 18940862.
- 54. Lucero DE, Ribera W, Pizarro JC, Plaza C, Gordon LW, Peña 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. https://doi.org/10.1371/journal.pntd.0003365 PMID: 25474154



- Karlsson AO, Holmlund G. Identification of mammal species using species-specific DNA pyrosequencing. Forensic Science International. 2007; 173(1):16–20. https://doi.org/10.1016/j.forsciint.2007.01.
 019 PMID: 17331687
- 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. https://doi.org/10.1007/s00414-006-0113-y PMID: 16845543
- 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. https://doi.org/10.1111/j.1556-4029.2007.00553.x PMID: 17868265
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990; 215(3):403–10. https://doi.org/10.1016/S0022-2836(05)80360-2 PMID: 2231712
- 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. PMID: 22499421
- Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. Electrophoresis 2010; 31(11):1764–72. https://doi.org/10.1002/elps.201000027
 PMID: 20446295
- 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. PMID: 3688448
- 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 cyto-chrome b heteroduplex-assay and sequencing for identification of triatomine blood meals. Infection, Genetics and Evolution. 2012; 12(1):21–7. https://doi.org/10.1016/j.meegid.2011.09.005 PMID: 21963963
- 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. https://doi.org/10.1371/journal.pntd.0001884 PMID: 23166846
- 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.
- 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. https://doi.org/10.1016/j.actatropica.2015.05.029 PMID: 26051910
- 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.
- 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. PMID: 9660458
- 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. PMID: 1801342
- 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. PMID: 8702018
- 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. PMID: 24180119



- **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.
- 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).