Sylvatic host associations of Triatominae and implications for Chagas disease reservoirs: a comprehensive review and new host records based on archival specimens

Anna Y Georgieva, Eric R L Gordon, Christiane Weirauch

1 Entomology, University of California, Riverside, Riverside, California, United States

Corresponding Author: Eric R L Gordon
Email address: egord003@ucr.edu

Background: The 148 species of kissing bug include important vectors of the debilitating, chronic, and often fatal Chagas disease, which affects several million people in Central and South America. An understanding of the natural hosts of this speciose group of blood-feeding insects has and will continue to aid ongoing efforts to impede the spread of Chagas disease. However, information on kissing bug biology is piecemeal and scattered, developed using methods with varying levels of accuracy over more than 100 years. Existing host records are heavily biased towards well-studied primary vector species and are derived from primarily three different types of observations, associational, immunological or DNA-based, with varying reliability.

Methods: We here gather a comprehensive and unparalleled number of sources reporting host associations via rigorous targeted searches of publication databases to review all known natural, or sylvatic, host records including information on how that record was collected. We integrate this information with novel host records obtained via attempted amplification and sequencing of a ~160 base pair (bp) region of the vertebrate 12S mitochondrial gene from the gastrointestinal tract of 64 archival specimens of Triatominae representing 19 species collected primarily in sylvatic habitats throughout the southern U.S. and Central and South America during the past 10 years. We show the utility of this method for uncovering novel and under-studied groups of Triatominae hosts, as well as detecting the presence of the Chagas disease pathogen via Polymerase Chain Reaction (PCR) of a ~400 bp sequence of the trypanosome 18S gene.

Results: New host associations for several groups of arboreal mammals were determined including sloths, New World monkeys, coatis, arboreal porcupines and, for the first time as a host of any Triatominae, tayras. A thorough review of previously documented sylvatic hosts, organized by triatomine species and the type of observation (associational, antibody-based, or DNA-based), is presented in a phylogenetic context and highlights large gaps in our knowledge of Triatominae biology.

Conclusion: The application of DNA-based methods of host identification towards additional species of Triatominae, including rarely collected species that may require use of archival specimens, is the most efficient and promising way to resolve recognized shortfalls.
Sylvatic host associations of Triatominae and implications for Chagas disease reservoirs: a comprehensive review and new host records based on archival specimens

Short title: Sylvatic hosts of Triatominae

Anna Georgieva*^1, Eric Gordon*^1, Christiane Weirauch1

1Department of Entomology, University of California Riverside, 900 University Avenue, Riverside, CA 92521, USA

* equal contribution; ^ corresponding author: egord003@ucr.edu

Abstract

Background:

The 148 species of kissing bug include important vectors of the debilitating, chronic, and often fatal Chagas disease, which affects several million people in Central and South America. An understanding of the natural hosts of this speciose group of blood-feeding insects has and will continue to aid ongoing efforts to impede the spread of Chagas disease. However, information on kissing bug biology is piecemeal and scattered, developed using methods with varying levels of accuracy over more than 100 years. Existing host records are heavily biased towards well-studied primary vector species and are derived from primarily three different types of observations, associational, immunological or DNA-based, with varying reliability.

Methods:
We here gather a comprehensive and unparalleled number of sources reporting host associations via rigorous targeted searches of publication databases to review all known natural, or sylvatic, host records including information on how that record was collected. We integrate this information with novel host records obtained via attempted amplification and sequencing of a ~160 base pair (bp) region of the vertebrate 12S mitochondrial gene from the gastrointestinal tract of 64 archival specimens of Triatominae representing 19 species collected primarily in sylvatic habitats throughout the southern U.S. and Central and South America during the past 10 years. We show the utility of this method for uncovering novel and under-studied groups of Triatominae hosts, as well as detecting the presence of the Chagas disease pathogen via Polymerase Chain Reaction (PCR) of a ~400 bp sequence of the trypanosome 18S gene.

Findings:

New host associations for several groups of arboreal mammals were determined including sloths, New World monkeys, coatis, arboreal porcupines and, for the first time as a host of any Triatominae, tayras. A thorough review of previously documented sylvatic hosts, organized by triatomine species and the type of observation (associational, antibody-based, or DNA-based), is presented in a phylogenetic context and highlights large gaps in our knowledge of Triatominae biology.

Conclusion:

The application of DNA-based methods of host identification towards additional species of Triatominae, including rarely collected species that may require use of archival specimens, is the most efficient and promising way to resolve recognized shortfalls.
Introduction

Triatominae, or kissing bugs, are blood-feeding members of the primarily predatory insect family Reduviidae (Order Hemiptera). This subfamily consists of 148 described species, nearly all distributed in the Americas, though several species occur in the Oriental region and *Triatoma rubrofasciata* (De Geer) is considered invasive throughout the tropics [1,2]. These insects are the sole vectors of *Trypanosoma cruzi* Chagas, the parasite responsible for Chagas disease. Chagas disease is a chronic debilitating disease, prevalent in Latin America, and affecting up to 10 million people worldwide [3]. There is no vaccine or effective cure once the symptoms of the chronic disease have manifested and the disease has been termed an emerging threat of the 21st century [4,5].

Most kissing bug species are suspected to be oligo- or polyphagous across a broad range of wild mammal and other vertebrate species [6,7]. Many Triatominae hosts are sylvatic mammals, but domestic mammals such as dogs, cats, and rodents can also be fed upon and act as reservoir hosts of the parasite [8]. Generalist kissing bug species also tend to feed on humans, e.g., *Triatoma infestans* Klug and *Rhodnius robustus* Larrousse, that can exhibit infection rates with *T. cruzi* higher than 40% [9,10]. Some triatomine species will target certain hosts if available and avoid other potential blood meal sources [1]. *Cavernicola pilosa* Barber appears to only feed on bats [11,12], while *Triatoma delpontei* Romana and Abalos and *Psammolestes* Bergroth species are usually found in association with various birds [11,13]. In addition, there are reports of some kissing bug species feeding on other arthropods [14–16] or exhibiting cleptohemophagy by feeding on other engorged kissing bug individuals [17]. The extent of both of these behaviors in a natural environment and for the great majority of kissing
bug species is unknown and may be driven by host availability. A diet lacking blood has been shown experimentally to result in complete mortality in at least some species [18], suggesting that arthropod feeding may be rare or driven by the lack of more suitable hosts. Overall, existing host association data is biased towards a handful of heavily studied and well-documented primary vector species and little data exists for many other Triatominae, particularly in sylvatic habitats [19]. Understanding patterns of host associations across Triatominae may help to elucidate their natural history and identify as yet underappreciated species of medical interest.

Records of Triatominae host associations are based on different types of observations that we here classify as associational (i.e. visual observation of a kissing bug in presumed or actual close association with a possible host), immunological and DNA-based methods. Each of these approaches possess a unique set of advantages and disadvantages. Many early studies on kissing bug-host interactions depended on observations of the insects during laboratory feeding tests or, more rarely, in the wild [11,20]. Despite presenting direct evidence for feeding, laboratory experiments are necessarily unnatural. Laboratory tests either offer insects a single possible host species and observe whether feeding occurs [21–23] or present them with a choice between two or more hosts to determine preference [24]. Both approaches may drive insects to feed on organisms that they would not feed upon in natural habitats. Observations of cohabitation of Triatominae with vertebrates in the wild, e.g., in animal nests and burrows, while more natural, are usually tentative because the association may not reflect actual feeding. There is likely also a bias towards more accessible terrestrial nests and burrows compared to corresponding arboreal habitats. Immunological methods to detect host associations in Triatominae were first used in the 1960s [25–27], seemingly overcoming problems posed by associational methods. By utilizing specific, but often polyclonal, antibodies in antisera developed for a predetermined range of
potential host species, experimenters infer direct host associations. Precipitin tests have continued to be a popular way to detect host antigens in Triatominae blood meals [28–30]. While a promising technique for forensic determination of actual hosts, disadvantages of precipitin tests are twofold. First, tests may suffer from non-specificity in antibody binding or irreproducibility as a result of variance in antibodies [31]. Second, due to the cost of developing different sets of antibodies for different groups of hosts, antibodies are typically developed to be specific only for large groups of potential host vertebrates, such as rodents. This approach results in reduced resolution of hosts, i.e. vertebrates are not identified to genus or species, but will also fail to detect unexpected or rare host taxa for which no antibody set is available.

More recently, studies have begun to use DNA-based methods to detect host identity [16,32,33]. These methods typically use PCR that targets the conserved regions of variable, mitochondrial genes for amplification. When amplified sequences are compared to a known database, it is usually possible to determine what species or at least genus of organism that the blood originated from. These studies have documented that certain species, such as *Triatoma rubida* (Uhler), *Triatoma protracta* (Uhler), and *Triatoma gerstaerckeri* (Stål), feed on a large variety of vertebrate hosts [16,34]. PCR-based studies also have the potential to determine the percentage of specimens within a given kissing bug population that have fed on humans [34]. While PCR is useful in detecting a wide range of hosts, given that databases such as GenBank now hold a library of barcodes for most mammal species and many other vertebrates, it does have a relatively high risk for human contamination [35]. Primers can also have biases in amplifying DNA that closely mirrors their sequence while not amplifying other sequences or also amplify the insect’s own DNA, which can interfere with detecting host DNA from the blood
sample. Multiple blood meals per specimen can amplify and interfere with determining a single
sequence and must be separated via cloning of the PCR product or next generation sequencing.

Despite the importance of identifying reservoir hosts of Chagas disease and the biology
of its vectors, most studies have focused on known primary vector species. These studies most
often targeted domestic or peridomestic habitats where the transmission risk to humans is
considered higher (e.g., [36,37]). In addition, most previous DNA-based studies have surveyed
narrow geographic areas, with several focusing on North America, and have used only live or
very recently preserved specimens. A comprehensive overview of triatomine-host associations
that specifies the method through which the record was obtained and allows for assessing the
reliability of the record is yet unavailable. Our study therefore has three objectives: 1) to
contribute to the growing knowledge base of host associations across Triatominae, we conducted
PCR of gastrointestinal contents extracted from 19 species of Triatominae from a range of
localities (Bolivia to U.S.); most specimens were collected in sylvatic habitats using light traps
and the sample comprises rarely encountered kissing bug species; 2) to determine the feasibility
of assessing host associations and trypanosome infection for archival kissing bug specimens that
have been preserved in ethanol for up to 10 years, we conducted PCR-based identification of
host sequences and trypanosomes for 64 specimens; 3) to establish currently documented
sylvatic host associations, we conducted a thorough literature review for all species of
Triatominae while recording the method used to determine that association; host patterns and
gaps in our current knowledge were visualized in phylogenetic context for both kissing bug
species and vertebrate hosts.

**Materials and Methods**
**Taxon sampling:** Triatominae specimens were primarily collected in sylvatic conditions via light trapping or hand collection throughout the southern U.S. and Central and South America and were preserved in ethanol (concentrations either unknown or 95%) between 2005 and 2015. We classify the habitat that each specimen as either domestic (found in a residence), peridomestic (found outside in a residential area, near a residence) or sylvatic (found in natural habitats, sometimes near field research stations). It was our aim to survey as wide a variety of Triatominae species as possible, but certain species (e.g., *Triatoma protracta*, *Panstrongylus geniculatus* [Latreille]) were sampled more thoroughly due to the availability of specimens present in the Weirauch lab ethanol repository. Voucher specimen data (unique specimen identifier [USI], determination, sex, specimen depository, collecting locality and event) were recorded using the Planetary Biodiversity Institute instance of the Arthropod Easy Capture database ([research.amnh.org/pbi/locality](https://research.amnh.org/pbi/locality)). Images of voucher specimens were taken using a Leica DFC450 C Microsystems system with a Planapo 1.0x objective. Voucher data, including collecting technique and images, are available online at Heteroptera Species Pages ([research.amnh.org/pbi/heteropteraspeciespage](https://research.amnh.org/pbi/heteropteraspeciespage)) and are best searched by species and then USI number.

**DNA extraction:** DNA of the gastrointestinal contents of Triatominae specimens was extracted in order to perform PCR. To avoid cross-contamination as well as the recognized threat of contamination with human DNA, all equipment and work benches were sterilized (dissecting petri dish, forceps, iris scissors) before and after processing each specimen using 10% bleach. Cuticular surfaces of specimens were sterilized with 1% bleach for 3 minutes, to eliminate possible contaminants acquired before capture in ethanol or during ethanol storage. The thorax and abdomen were separated and contents of the abdomen were removed with forceps and
placed into an Eppendorf tube. While performing this procedure, the sex of the specimen and
whether blood was visible in the gut or not was recorded. When a large volume of blood was
present, contents were divided into separate Eppendorf tubes. Gut contents were homogenized
for 2 minutes with an Eppendorf pestle and DNA was extracted with a QIAGEN DNeasy blood
and tissue kit. We recorded the amount of blood in each specimen on a scale of 1 to 4 where: 1 --
no material visible in digestive tract; 2 -- small amount of dark, digested blood present; 3 --
obvious blood present; 4 -- completely engorged with blood. For seven specimens of *Triatoma*
*protracta* and one specimen of *Eratyrus mucronatus* Stål which had been extracted previously,
we were not able to record the amount of blood present in the digestive tract.

**PCR:** We tested seven previously developed sets of primers targeting three different
mitochondrial genes (Table 1) designed for identification of vertebrate hosts from invertebrate
blood meals via PCR on each of our extracts. We were not able to achieve consistent, acceptably
broad or specific results for several sets of previously used primers listed in Table 1 (all those
listed without asterisks). These primers amplified the corresponding DNA sequence of certain
species of Triatominae or did not amplify DNA that should be present based on the results of
other primer sets. We found that the “Kitano” 12S primers [38] yielded the most consistent and
the greatest number of amplified bands across samples among the primer sets tested and this is
the only primer set for which we present results. If samples yielded multiple bands after
electrophoresis on 1% agarose gels, they were gel extracted using QIAquick Gel Extraction kit.
If samples extracted from insects containing a large volume of blood did not result in bands after
PCR, the extract was diluted 1 in 10 and PCR was conducted again to ensure that a PCR-
inhibitor from the blood meal was not present (e.g., approach successful with sample UCR_ENT
00123869). We used primers for the 18S region to determine presence of trypanosomes in our
extracts (Table 1; [39]). PCR conditions consisted of an initial denaturation step of 94˚C for 5 minutes, denaturation at 94˚C for 30 seconds, the annealing temperature listed in Table 1 for 30 seconds, extension at 72˚C for 30 seconds repeated for 35 cycles with a final extension time at 72˚C for 10 minutes.

**Table 1. Primer sequences and PCR conditions used in this study.**

<table>
<thead>
<tr>
<th>DNA target</th>
<th>Primer set</th>
<th>Locus</th>
<th>Direction</th>
<th>Sequence</th>
<th>Annealing temperature</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertebrate DNA</td>
<td>Kitano 12S*</td>
<td>12S</td>
<td>F</td>
<td>5'-CCC AAA CTG GGA TTA GAT ACC C-3'</td>
<td>57˚</td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>5'-GTT TGC TGA AGA TGG CGG TA-3'</td>
<td>53˚</td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td>Melton 12S</td>
<td>12S</td>
<td>F</td>
<td>5'-ACT GGG ATT AGA TAC CCC ACT ATG-3'</td>
<td>53˚</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>5'-ATC GAT TAT AGA ACA GGC TCC TC-3'</td>
<td>45˚</td>
<td>[41]</td>
</tr>
<tr>
<td>Vert COI</td>
<td>M13BC-FW</td>
<td>COI</td>
<td></td>
<td>5'-TGT AAA ACG ACG GCC AGT HAA YCA YAA RGA YAT YGG NAC-3'</td>
<td>45˚</td>
<td>[42]</td>
</tr>
<tr>
<td>DC-CytB</td>
<td></td>
<td>UP</td>
<td></td>
<td>5'-CRT GAG GGC AAA TAT CHT TYT-3</td>
<td>42.5˚</td>
<td>[43]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DW</td>
<td></td>
<td>5'-ART ATC ATT CGW GTT TAA TRT-3</td>
<td>42.5˚</td>
<td>[43]</td>
</tr>
<tr>
<td>Avian CytB</td>
<td>CytB</td>
<td>F</td>
<td></td>
<td>5'-GAC TGT GAC AAA ATC CCN TTC CA-3'</td>
<td>55˚</td>
<td>[44]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td></td>
<td>5'-GGT CTT CAT CTY HGG YTT ACA AGA G-3'</td>
<td>55˚</td>
<td>[44]</td>
</tr>
<tr>
<td>Mammalian CytB</td>
<td>CytB</td>
<td>F</td>
<td></td>
<td>5'-CGA AGC TTG ATA TGA AAA ACC ATC GTT G-3'</td>
<td>55˚</td>
<td>[44]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td></td>
<td>5'-TGT AGT CRT CWG CTT CTA-3'</td>
<td>55˚</td>
<td>[44]</td>
</tr>
<tr>
<td>Vert CytB</td>
<td>CytB</td>
<td>CB1-L</td>
<td></td>
<td>5'-CCC CTC AGA ATA TTG CTC CTY A-3'</td>
<td>55˚</td>
<td>[45]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CB2-H</td>
<td></td>
<td>5'-CAT CCA ACA TCT CAG CAT GAA A-3'</td>
<td>55˚</td>
<td>[45]</td>
</tr>
<tr>
<td>Trypanosome DNA</td>
<td>Tez*</td>
<td>18S</td>
<td>18sf</td>
<td>5'-TTA ACG GGA ATA TCC TCA GC-3'</td>
<td>50˚</td>
<td>[46]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S829r</td>
<td>5'-GCA TCA CAG ACC TCG TG TGT-3'</td>
<td>50˚</td>
<td>[47]</td>
</tr>
</tbody>
</table>

**Purification, sequencing and analyzing:** All PCR products were cleaned using SureClean (Bioline) before sequencing using the Macrogen EZ-Seq service. Once sequences were obtained, the program Sequencher was used to process chromatographs. Blastn was then used to compare sequences to the GenBank database. Sequences of the 12S gene were considered to be derived from the same species as the closest match represented in GenBank only if they were 100% identical and most or all other members of that genus were also represented and differed in sequence (e.g. sequences KX779919, 100% to Choloepus didactylus, and KX779929, 100% to
Neotoma lepida) otherwise we classified it only to genus (e.g., sequence KX779923, 100% to 
Lagotricha lagotricha) or to an even higher level (e.g., sequence KX779938, 100% to Mustela 
kathiah and classified as Mustelidae sp.). When multiple members of a genus were represented 
in GenBank, but none matched 100%, we classified our sequence to that genus if it was more 
than 98% identical to one member of the genus and closer to other members of that genus than to 
any other genus (most sequences, e.g. sequences KX779920 98.6% to Dasyprocta leporina and 
KX779934 98.7% to Saguinus oedipus). Trypanosome-derived PCR products of the 18S rRNA 
gene were sequenced and the sequence compared to the GenBank database for identification.

Phylogeny construction: We gathered all available data for Triatominae species and 
closely related reduviids (Stenopodainae, Zelurus spp. Opisthacidius spp.) on GenBank totaling 
9,343 bp from 120 taxa as aligned using the MAFFT EINS-i algorithm [48] comprising the loci 
from the nuclear rRNA operon and the mitochondrial genome (18S rRNA, ITS 1 5.8 rRNA and 
ITS 2, D2, D3-D5 regions of the 28S rRNA; 16S rRNA COI, COII, Cytb) and constructed a 
phylogeny using a partitioned RaxML analysis (Fig. S1, sampled species in red) partitioned by 
the best scheme as determined by PartitionFinder (Phylip, Data S1; 9 partitions, Data S2). We 
excluded Belminus herreri Lent & Wygodzinsky from our final analysis due to the 
reconstruction of this taxon as sister to T. rubrofasciata with low support and a long terminal 
branch; this result may be due to this taxon only being represented by the 18S gene that has low 
phylogenetic utility for relatively recent divergences and the lack of that gene in the only other 
representative of the tribe Bolboderini in our dataset, Microtriatoma trinidadensis (Lent) (rRNA 
28S D2 region only). To visualize all species of Triatominae in a single figure, we placed all taxa 
for which molecular data are currently unavailable to this backbone phylogeny using information 
on morphological similarities from [6].
Literature review: We used a Web of Science search (all databases) to query the species name of all the 148 currently recognized Triatominae species and surveyed results for literature records of host associations. For all matches, we assessed titles and abstracts for the inclusion of information on host associations and if we determined that they may include relevant information, we scrutinized the publication for the record of the host association and the type of method with which it was achieved. For a few widely studied organisms with an extremely high number of matching publications (e.g., *Triatoma infestans*, *Rhodnius prolixus*), we limited our searches with additional target words such as “host” or “association” to increase the feasibility of surveying all relevant results. Every attempt was made to find the primary source of a host record but occasionally, we resorted to using a source that also included a review of other sources and for which the source of the record was unclear and may be either primary or secondary. We limited the inclusion of host records to evolutionarily relevant host species by excluding laboratory results or host associations of exclusively domestic animals such as dogs, chickens or other farmed animals due to the unnatural presence of these hosts. An exception was made for domestic rodents, including mice and rats, that also occur in natural environments. Host taxa were divided into major groups with all arthropods, amphibians, birds and reptiles each comprising a single group and the remaining mammals split primarily by order with some exceptions (suborder level for xenarthrans, superfamily level for primates and rodents, family level for carnivores).

Results and Discussion

Of 64 total specimens tested (38 males and 26 females), we were able to determine a host association for 24 specimens (37.5%) with a maximum of a single host determined per specimen. Of hosts with observable blood (28 out of 56 with observations recorded; quantity of blood
categories 2-4), 18 or 64.3% gave positive host results, compared to 2 or 7.1% of those observed without visible blood (category 1) in the digestive tract for which we obtained a host sequence. Specimens with observable blood but without amplifiable host DNA may represent samples with DNA degraded beyond allowing for amplification with primers targeting the ~160 bp region of the 12S gene. Alternatively, it is possible that the Kitano 12S primers do not amplify 12S sequences from certain hosts, though they were designed based on sequences from divergent vertebrate sequences from sharks to fish to reptiles to amphibians and mammals (Kitano 2007). Of the 24 specimens with host determinations, 17 or 70.8% were male, a heavily skewed ratio. Similarly, of the 28 specimens with visible blood in the digestive tract, 21 or 75% were male. We speculate that blood-fed males may be more predisposed to dispersing in flight while searching for mates and thus be more susceptible to light traps. In contrast, females may be stationary after feeding, attempting to find a suitable place for oviposition, possibly near or in the same location as her blood meal. The infection rate of all specimens with *T. cruzi* across sampled Triatominae was 31.3% or 20 specimens with one specimen of *Rhodnius pictipes* producing a band that, after sequenced, matched that of *Trypanosoma rangeli* Tejera. Of the 20 specimens testing positive for *T. cruzi*, 50% also possessed detectable host DNA and 50% did not. *Panstrongylus geniculatus* was the species with the highest number of *T. cruzi* positive individuals (5/11) followed by *Triatoma protracta* (4/17) and all three individuals tested of *T. dimidiata* which was the species with the highest positive percentage rate (100%; 3/3) along with *T. dispar* (100%; 1/1) and *T. recurva* (100%; 1/1). Specimens of seven additional species also tested positive for *T. cruzi* (Table 2), all of which were previously known to be capable of hosting the parasite [7].
### Table 2. Specimen data including data obtained by vertebrate and trypanosome

<table>
<thead>
<tr>
<th>USI</th>
<th>species</th>
<th>Country: Primary collection</th>
<th>Location</th>
<th>Date of Collection</th>
<th>Host (128 sequence)</th>
<th>Inferred Host</th>
<th>BLAST host</th>
<th>BLAST hit %</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCR_ENT1</td>
<td>Acanthocheilus robustus</td>
<td>NICARAGUA; Heredia</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT2</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT3</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT4</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT5</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT6</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT7</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT8</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT9</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT10</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT11</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR_ENT12</td>
<td>Panstrongylus rufotuberculatus</td>
<td>NICARAGUA; Cayenne</td>
<td>2007</td>
<td>95%</td>
<td>M</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DNA targeted PCR.**
always resulted in a single, uncontroversial host sequence, similar to results from other studies conducted using similar primers [49], but in contrast to other studies where cloning of the PCR product was performed and as many as four hosts were detected from a single specimen [50].

The archival nature of our specimens may have contributed towards the lack of detection of blood meals other than the one with most abundant DNA available. The 24 hosts detected
represented mainly animals present in a sylvatic environment with the exception of three records of the domestic dog from *T. protracta* (*n* = 17) specimens collected in Southern California in residential areas, one record of a deer mouse detected from *T. pallidipennis* (*n* = 1) at a residence in Mexico and two records of humans from sylvatic areas in Ecuador and Bolivia from *Rhodnius pictipes* (*n* = 4) and *Panstrongylus rufotuberculatus* (*n* = 3), respectively. These results represent fewer human records than some similar DNA-based studies (48.8% in [50]; 38% in [34]), which may reflect the sylvatic nature of most of our collecting sites and, potentially, the rigorous sterilization protocol we utilized. Because we did not obtain any sequences from marsupials or armadillos, it is possible that the Kitano 12S primers do not amplify 12S sequences from these hosts. While the primers match the corresponding sequence from armadillos at all but two mismatches at beginning of the reverse primer, the sequence in opossums appears to have several mismatches with the reverse primer that may have prevented binding and amplification. Of the sylvatic hosts detected, many are arboreal mammals and a surprising number reflect new host species, or new records for the larger groups of vertebrates they belong to (Fig. 1, dark red rectangles) as defined in Figure 1. Even after an extensive literature review, we were unable to find any record of these new groups of animals being recorded for hosts of these species. The new hosts detected for *Panstrongylus geniculatus* (*n* = 11) include sloths (*Choloepus didactylus* Linnaeus), the arboreal weasel-like animal known as a tayra (*Eira barbara* [Linnaeus]), an unidentified member of Mustelidae, New World monkeys (*Lagothrix* sp. and *Saguinus* sp.) and agoutis (*Dasyprocta* sp.), all but the last, representing the first record for that group of vertebrate for that species. The tayra record represents the first record of any Triatominae species feeding on this species. We recovered several new host records for New World monkeys, apart from the two species identified for *P. geniculatus*, including *Cebus* sp. And *Saimiri* sp., *Ateles* sp., and
Cebus sp. For Rhodnius robustus (n=7), Panstrongylus rufotuberculatus (n=3) and T. dimidiata (n=3), respectively. These all represent the first New World monkey hosts recorded for those species except for T. dimidiata which has been previously noted as possessing antigens reacting to antibodies developed for detecting New World monkey specific proteins. Our new host records also included the arboreal porcupine genus Coendou for P. rufotuberculatus, coati (Nasua sp.) for R. prolixus (n=1) and a sequence with the closest match to a kinkajou (Potos sp.) but only determined to the level of Procyonidae for Triatoma dispar Lent (n=1), all new animal groups for those species. The prevalence of arboreal mammals among our new host records may reflect the inaccessibility of these habitats for gathering associational observations or absence of antigens targeting these taxa in previous studies. Potentially, these under recognized groups of hosts play a more prominent role in sylvatic cycle of T. cruzi than has been previously understood.
Figure 1. A visualization of the known sylvatic hosts of Triatominae and the type of record(s) supporting that association. The animals at the top of the figure represent the following groups from left to right, alternating from top to bottom: arthropod (Arthropoda), amphibian (Amphibia), lizards (Lepidosauromorpha), bird (Aves), opossums (Didelphimorphia), armadillo (Cingulata), sloth (Folivora), anteater (Vermilingua), shrew (Eulipotyphla), bat (Chiroptera), even-toed ungulate (Artiodactyls), feline (Felidae), canine (Canidae), musteloid (Mephitidae; Skunk), musteloid (Procyonidae; Raccoons and relatives), musteloid (Mustelidae; Weasel), platyrhine monkey (Ceboidea), human (Hominidae), rabbits (Lagomorpha), rodent (Sciuromorpha), rodent (Muroidea), rodent (Geomyoidea), rodent (Octodontoidea), rodent (Chinchilloidea), rodent (Cavioidae), rodent (Erethizontoidea). Relationships among mammals were simplified from [51] for Carnivora, [52] for deep level relationships and [53] for rodents. Triatominae taxa in light grey and indicated with an asterisk were added to the phylogeny based on morphological similarities indicated in the literature. Filled in matrix rectangle indicate observations with green indicating antibody-based observations; blue, associational; and red, DNA-based with bright red indicating newly obtained DNA-based host records for that species in our study. Colored outlines of rectangles across all represented rodent superfamilies represent observations specified only to the order Rodentia.
As a result of our literature review, we were able to detect some patterns that have not been previously widely recognized though it should be noted that host preferences cannot be determined with this data and a lack of evidence of a particular record does not indicate that such a host association may not be uncovered in the future. All of our aggregated data is reported in Table S1 with full references in Article S1. The number of species with associations with amphibians (11), reptiles (30) and birds (46, or 32% of all Triatominae species) is higher than we would have expected as these species tend to be thought of as minor hosts for Triatominae or only primarily associated with certain species [54]. This discrepancy may reflect that these hosts are not known to be able to harbor the *T. cruzi* pathogen [55,56] and deliberate attention has been paid to mammalian reservoir hosts. The most common group of known hosts among all species of Triatominae are marsupials (49 species), birds (46 species) and Muroidea including rats and mice (45 species). Humans rank as the fourth most commonly associated group with species of Triatominae (40 species) although it is possible that some human records have evaded our attention. Some of the rarest groups of known sylvatic hosts are shrews, artiodactyls and felines known only from a single host record, although this excludes records associated only with domestic species of those groups such as the domestic cat and the domestic pig. The rodent group Chinchilloidea which includes chinchillas had surprisingly few records (only for *Mepraia spinola* [Porter] and *Triatoma infestans*, both associated with viscachas in the genus *Lagidium*).

There are 51 kissing bug species for which we were unable to find any records for sylvatic hosts. The largest group with scarce host data are the Old World species including the genus *Linshcosteus*, where associations with rats, birds, and humans have been recorded for *T. rubrofasciata* and only humans for *Triatoma cavernicola* Else & Cheong. Species which have been studied using DNA-based methods tend to have a broader range of known hosts than...
species only investigated using associational and immunological approaches. For example, *Rhodnius pallescens* Barber and *T. gerstaeckeri* have the broadest ranges of recorded host groups (20 and 15 respectively, out of 26 groups recorded across all species) primarily because of single studies focused on each of those species [49,57]. Some species thought to have narrow host ranges show some evidence, although mostly associational and rarely antibody-based, of also feeding on additional hosts such as rodents for *Cavernicola*, normally thought to associate only with bats, and both rodents and marsupials for *Psammolestes*, normally associated with birds.

There is no obvious host specificity of Triatominae clades and, overall, our review points towards the generalist bent of all Triatominae species, perhaps influenced more by hosts present in a habitat than innate preference towards certain groups of hosts.

**Conclusion**

We believe that DNA-based methods for determining host associations of blood-feeding species offers the best route forward in understanding the biology and epidemiological importance of these speciose group of vectors. While there may be bias in representation of known vertebrates on public databases, these databases continue to encompass data for an ever growing range of species. We recommend the deposition of sequences acquired from vertebrate hosts of vectors into public databases even if their identity is not known at the time, not yet routinely done, as the influx of known sequences may later shed light on their identity. While DNA-based studies can suffer from primer specificity biases, rigorous testing of primers sets for universality and specificity can minimize this possibility. We have shown that this method can be used on archival specimens and we recommend future efforts using this relatively cheap, efficient and effective method in order to better understand the habits of all species of these vectors, particularly those that are less well studied. We found that 95% ethanol seemed to
preserve DNA well enough for amplification of the 150 bp chunk of 12S rRNA using our preferred set of primers (38). While we did achieve some results for specimens without observable blood in the digestive tract, we had a much higher success rate for those with observable blood and time and energy could be saved by not extracting DNA from the former group of specimens.

Acknowledgements

We thank Jason Cryan, Andrew Ernst, Dimitri Forero, Sarah Frankenberg, Jeremy Huff, Warren Macdonald, Michael E. Irwin, J.J. Ramirez, Gavin Svenson, Julie Urban, and Michael Whiting for donations of specimens.


40. Alcaide M, Rico C, Ruiz S, Soriguer R, Muñoz J, Figuerola J. Disentangling vector-borne transmission networks: a universal DNA barcoding method to identify vertebrate hosts from...


57. Gorchakov R, Trosclair LP, Wozniak EJ, Feria PT, Garcia MN, Gunter SM, et al. Trypanosoma cruzi infection prevalence and bloodmeal analysis in triatomine vectors of Chagas disease from rural
doi:10.1093/jme/tjw040

Supplemental documents

Fig. S1. RaxML tree of Triatominae species with molecular data available (excluding Belminus herreri) with bootstrap support values shown. Kissing bug species in red are those for which gut extracts were assayed with primers for vertebrate host and trypanosome DNA.

Table S1. Table of sylvatic host records of all Triatominae species.

Article S1. References for Table S1

Data S1. Phylip file of alignment

Data S2. Partitioning scheme as determined by PartitionFinder.