

Novel polymerase chain reaction-restriction fragment length polymorphism assay to determine internal transcribed spacer-2 group in the Chagas disease vector, *Triatoma dimidiata* (Latreille, 1811)

Bethany Richards¹, Nicholas M de la Rúa², Carlota Monroy³, Lori Stevens², Patricia L Dorn^{1/+}

¹Loyola University New Orleans, New Orleans, LA, USA ²University of Vermont, Burlington, VT, USA

³Universidad de San Carlos, Guatemala City, Guatemala

Triatoma dimidiata is the most important Chagas disease insect vector in Central America as this species is primarily responsible for *Trypanosoma cruzi* transmission to humans, the protozoan parasite that causes Chagas disease. *T. dimidiata* sensu lato is a genetically diverse assemblage of taxa and effective vector control requires a clear understanding of the geographic distribution and epidemiological importance of its taxa. The nuclear ribosomal internal transcribed spacer 2 (ITS-2) is frequently used to infer the systematics of triatomines. However, oftentimes amplification and sequencing of ITS-2 fails, likely due to both the large polymerase chain reaction (PCR) product and polymerase slippage near the 5' end. To overcome these challenges we have designed new primers that amplify only the 3'-most 200 base pairs of ITS-2. This region distinguishes the ITS-2 group for 100% of known *T. dimidiata* haplotypes. Furthermore, we have developed a PCR-restriction fragment length polymorphism (RFLP) approach to determine the ITS-2 group, greatly reducing, but not eliminating, the number of amplified products that need to be sequenced. Although there are limitations with this new PCR-RFLP approach, its use will help with understanding the geographic distribution of *T. dimidiata* taxa and can facilitate other studies characterising the taxa, e.g. their ecology, evolution and epidemiological importance, thus improving vector control.

Key words: Chagas disease - Triatominae - RFLP - methods - population genetics

Importance of Chagas disease and Triatoma dimidiata taxa - *Trypanosoma cruzi*, the causative agent of Chagas disease, is most frequently transmitted to humans by the insect vector, *T. dimidiata*, in Central America where it shows remarkable genetic divergence across its geographic range. More than 25 million people in Latin America are at risk for Chagas disease and effective vector control and consequently prevention of transmission require clear taxonomy and understanding of the subdivisions of this genetically diverse vector. Taxa of *T. dimidiata* may differ in attributes that contribute to vector competence, such as habitat preference, feeding patterns, infection prevalence with *T. cruzi* etc. (Dorn et al. 2007). Clear taxonomy allows one to target control efforts to the most epidemiologically important populations. The DNA sequence of the nuclear ribosomal internal transcribed spacer 2 (ITS-2) is a frequently used marker to infer the systematic subdivisions of triatomines because it is presumed selectively neutral and has a relatively high mutation rate. Based on this and other markers, the taxonomy has been revised multiple times (Marcilla et al. 2001, Dorn et al. 2009). The large number of ITS-2 sequences available in public databases also increases its value as a genetic marker. However,

oftentimes amplification of ITS-2 is difficult due to its relatively long length and sequence composition (Mas-Coma & Bargues 2009). The high occurrence of insertions and deletions and the relatively low copy number of this nuclear marker in comparison to mitochondrial DNA markers tend to make amplification and alignment analysis of ITS-2 difficult in triatomines, especially at higher taxonomic levels (Mas-Coma & Bargues 2009). In order to successfully amplify and sequence ITS-2, we have developed a new primer set amplifying a shorter, less problematic fragment of ITS-2. Furthermore, we have developed a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assay for quickly assessing ITS-2 group identity which has reduced, but not eliminated, the requirement for DNA sequencing to correctly identify ITS-2 group.

Previous technique and disadvantages - Four genetically distinct taxa of *T. dimidiata* have been identified in southern Mexico, Central America and northern South America by ITS-2 (groups 1A, 1B, 2 and 3). Group 3 appears to be a cryptic species due to its genetic divergence and populations have been identified in Peten, Guatemala, Yucatan, Mexico, Yoro, Honduras, and Cayo, Belize (Dorn et al. 2007, 2009, Bargues et al. 2008). The primers most commonly used to determine ITS-2 taxa are complementary to the conserved, flanking regions of 5.8S and 28S rDNA and amplify a ~900 base pairs (bp) fragment containing the complete sequence of ITS-2 (Marcilla et al. 2001). However, with these primers, oftentimes the template fails to amplify and/or produce interpretable sequence. This is likely due to (i) the large size of the PCR product which may be compromised in

doi: 10.1590/0074-0276108042013001

Financial support: NIH (1R15 A1079672-01A), Richard Frank Grant, the Mullahy, SJ Funds/Loyola University

+ Corresponding author: dorn@loyno.edu

Received 21 August 2012

Accepted 26 March 2013

older and somewhat degraded DNA and (ii) DNA polymerase slippage due to the extensive homopolymeric and microsatellite regions present within the 5'-most ~100 bp of ITS-2. Group 3 specimens in particular have longer and more complicated microsatellite regions making amplification and sequencing more challenging (PL Dorn, unpublished observations). Many investigators have had higher success by sequencing cloned ITS-2 fragments. However, even when successful, cloning and sequencing are expensive, labour intensive and often not feasible for laboratories in endemic countries.

New technique and advantages - To overcome these challenges we redesigned the primers to amplify only the 3'-most 200 bp of ITS-2. Analysis of all 39 *T. dimidiata* ITS-2 sequences present in GenBank® (Benson et al. 2012) in July 2011 reveals that this region is sufficient to distinguish the ITS-2 group, bypassing the troublesome microsatellite and homopolymeric regions. Furthermore, we have developed a RFLP assay of the PCR product that determines ITS-2 group without sequencing for most samples. PCR-RFLP has been previously used to distinguish species within many insect vector complexes, for example, the malarial insect vector in the *Anopheles* complex (Van Bortel et al. 2000, Fanello et al. 2002). Using a similar method, PCR-RFLP can be used to distinguish ITS-2 groups of *T. dimidiata*. Our approach is a less expensive and technically easier method for determining the ITS-2 group of *T. dimidiata*, so is amenable to endemic countries and provides much needed genetic information more quickly.

MATERIALS AND METHODS

T. dimidiata specimen information and DNA isolation - All *T. dimidiata* (334 specimens) were collected by trained personnel across the geographic range and stored at -20°C in 95% ethanol/5% glycerol until DNA isolation. The *T. dimidiata* were transported to Loyola University New Orleans and the DNA was isolated from three-five legs of each specimen using the DNeasy Blood and Tissue kit (Qiagen, Valencia, CA) following the manufacturer's protocol.

Primer design and RFLP assessment - Thirty-nine haplotypes of *T. dimidiata* ITS-2 sequences falling into four distinct genetic "groups" have been identified (Dorn et al. 2007, 2009, Bargues et al. 2008). A microsatellite and homopolymeric region (identical or similar to (TA)₅T₃(TA)₇ and (T)₉) are present within the first 100 bp near the 5' end of all of these ITS-2 sequences. Previous studies revealed a longer microsatellite repeat, such as ((TA)₅T₃(TA)₈), in many specimens that failed to amplify and/or sequence properly, specifically specimens from group 3 (PL Dorn, unpublished observations). Therefore, we designed new primers using Primer 3 in MacVector (version 11.1.2, Cary, NC) avoiding the 5' end and amplifying a relatively small product that is still able to distinguish between all four genetic groups. The new forward primer (ITS2_200F) anneals at position 257 of ITS-2: 5'TCGYATCTAGGCATTGTCTG 3', Tm 67°C, GC 67% [degenerate to accommodate a single nucleotide polymorphism (SNP) in two haplotypes] and

the new reverse primer (ITS2_200R) anneals within the conserved 28S rDNA at position 69: 5' CTCGCAGC-TACTAAGGGAATCC 3', Tm 64°C, GC 55%. These result in an amplicon of ~320 bp, of which ~200 bp are ITS-2 and 91 bp are 28S rDNA, excluding the primers.

For 171 restriction enzymes, potential diagnostic restriction sites in the PCR product (i.e., within the last ~200 bp of ITS-2 amplified by the new primers) were assessed *in silico* using MacVector for all 39 published ITS-2 haplotypes GenBank® (Benson et al. 2012). An assay was developed using two enzymes, BsrGI followed by ApoI, which showed unique restriction sites for the ITS-2 groups.

PCR and sequencing - The ITS-2 region was amplified from 3% of the isolated DNA in a 30 µL reaction (3.5 mM MgCl₂, 2 U Taq DNA polymerase, 200 µM dNTPs, 10 pmol ITS2_200 primers). Amplification conditions were as follows: initial denaturation at 94°C 2 min, followed by 30 cycles of 94°C 30 sec, 58°C 30 sec, 72°C 30 sec and a final polishing step of 72°C for 7 min. Five microlitres of the PCR products were visualised by ultraviolet (UV) transillumination after 1% agarose gel electrophoresis and ethidium bromide staining. DNA of successful amplifications was recovered using QIAquick PCR Purification kit or the QIAquick Gel Extraction kit (QIAGEN, Valencia, CA), sequenced in both directions (Retrogen Inc, San Diego, CA), and sequences were edited and aligned using CLUSTALW 2.0 in MacVector (Thompson et al. 1994). ITS-2 groups of all trimmed and edited sequences were inferred using BLAST (Altschul et al. 1990) (blast.ncbi.nlm.nih.gov/Blast.cgi).

RFLP assay - After amplification with the ITS2_200 primers, restriction enzyme digests were performed on 12 specimens using all 100 µL of the amplified and purified PCR product following manufacturer's instructions (New England Biolabs, Inc, Ipswich, MA). Digested PCR products were visualised by UV transillumination after 2% Metaphor Agarose (Lonza, Rockland, ME) gel electrophoresis and ethidium bromide staining.

RESULTS

Using the primers designed by Marcilla et al. (2001), the 900 bp ITS-2 fragment was successfully amplified from only 59.6% (199/334) of samples by PCR. Using our newly developed ITS2_200 primers, we successfully amplified over 90% (78/86) of samples, of which 85% were successfully sequenced (66/78).

In silico analysis shows the PCR-RFLP assay using a sequential digest with the restriction enzyme BsrGI followed by ApoI successfully distinguishes all known ITS-2 group 3 haplotypes from groups 1A, 1B and 2 and distinguishes all of the group 2 haplotypes with exception of a rare group 2 variant. First, after amplification with the ITS2_200 primer set, digestion of the PCR product with BsrGI cleaves nine out of the 10 ITS-2 group 2 haplotypes into two fragments (193 and 125 bp) (Fig. 1). The other ITS-2 groups as well as a rare group 2 variant lack these sites so are left undigested. Second, following digestion with BsrGI, digestion with ApoI cleaves the undigested PCR product into two (groups 1A, 1B and the rare group 2 variant, 255 and 63 bp) or three fragments

(ITS-2 group 3, 151, 102 and 63 bp) (Fig. 1). Finally, if it is important, groups 1A and 1B (and the rare group 2 variant) can be identified by sequencing.

In order to further verify the *in silico* results, three *T. dimidiata* specimens per ITS-2 group (the group determined previously by sequencing) were tested with this PCR-RFLP assay. In all cases the observed RFLP banding patterns confirmed the expected results (Fig. 2, Table).

DISCUSSION

The use of the ITS2_200 primers to amplify a smaller (300 bp vs. 900 bp) region of ITS-2 leads to a substantially higher rate of success in amplification and sequencing, which provides previously unavailable valuable genetic information for specimens from geographic areas with high human infection rates. The ITS2_200 primers were found to be especially useful for group 3 specimens that previously were difficult to assess because of longer microsatellite sequences and for specimens with old and potentially degraded DNA. This higher rate of success could be due to the significant decrease in size of the amplified product, from ~900--~300 bp, as well as avoiding the problematic homopolymeric and microsatellite regions at the 5' end. These new primers are located in highly conserved areas of the ITS-2 and 28S rDNA. Results of a BLAST search (Altschul et al. 1990) indicate the primers can be used to amplify a similar product in other Triatominae (data not shown).

By using an RFLP assay, it is now possible to rapidly distinguish ITS-2 groups 2 and 3 from group 1 using basic PCR and gel electrophoresis techniques. This will be especially helpful in endemic areas where groups 2 and 3 specimens occur in sympatry (Dorn et al. 2009). Theoretically, this PCR-RFLP approach could distinguish groups 2/3 hybrids by producing overlapping group 2 and group 3 banding patterns, although this requires fur-

ther investigation (Herrera-Aguilar et al. 2009). Because ITS-2 is a multicopy cistron, there is a possibility that individual copies could contain SNPs at the restriction site. As rare variants these would likely be missed by examining the gels, as well as in the original sequencing protocol which reveals the predominant sequence. A careful review of the chromatograms of all specimens revealed no significant peak overlap at restriction sites, as would be expected if there was significant copy variation. To determine ITS-2 group with 100% accuracy, the DNA sequence of many fewer samples will need to be determined (those scored as 1A or 1B, which includes the rare group 2 variant, by the RFLP assay). Fortunately, previous data (Bargues et al. 2008, Dorn et al. 2009) show that 1A and 1B can usually be identified by geographic origin. Because the PCR-RFLP assay can be used to quickly determine ITS-2 information for *T. dimi-*

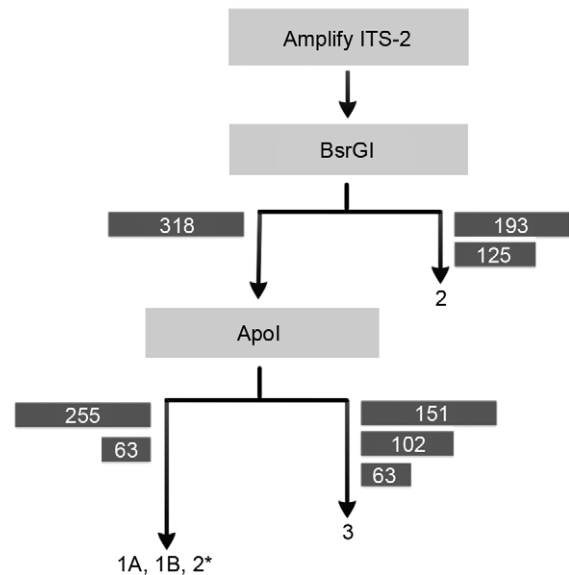


Fig. 1: restriction fragment length polymorphism assay for internal transcribed spacer 2 (ITS-2) groups of *Triatoma dimidiata*. This figure shows the unique fragment lengths produced for ITS-2 groups of *T. dimidiata* after digestion with enzymes: BsrGI and ApoI. Asterisk means one haplotype of ITS-2 group 2.

TABLE
Triatoma dimidiata collection
information and gel banding patterns

Gel lane	Origin	ITS-2 group determined by sequence	Gel banding pattern (bp)
1	Yucatan, Mexico	2	193, 125
2	Yucatan, Mexico	2	193, 125
3	Yucatan, Mexico	2	193, 125
4	Yucatan, Mexico	3	151, 102, 63
5	Yucatan, Mexico	3	151, 102, 63
6	Peten, Guatemala	3	151, 102, 63
7	Alta Verapaz, Guatemala	1A	255, 63
8	Alta Verapaz, Guatemala	1A	255, 63
9	Jutiapa, Guatemala	1A	255, 63
10	Santander, Colombia	1B	255, 63
11	Santander, Colombia	1B	255, 63
12	Santander, Colombia	1B	255, 63

bp: base pairs; ITS-2: internal transcribed spacer 2.

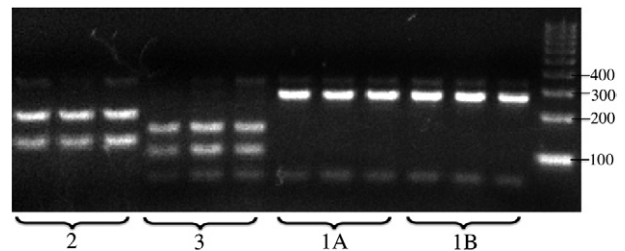


Fig. 2: gel banding patterns after restriction digests of *Triatoma dimidiata*. This figure shows the unique gel banding patterns of three *T. dimidiata* specimens from each internal transcribed spacer 2 group (shown below gel lanes) following digestion with enzymes BsrGI and ApoI.

diata in areas where sequencing might be unavailable or too expensive, its use will facilitate other studies in the ecology, epidemiological importance and geographic distribution of *T. dimidiata* taxa. The numerous applications of this method can be used to develop more targeted and effective vector control.

REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ 1990. Basic local alignment search tool. *J Mol Biol* 215: 403-410.
- Bargues MD, Klisiowicz DR, Gonzalez-Candelas F, Ramsey JM, Monroy C, Ponce C, Salazar-Schettino PM, Panzera F, Abad-Franch F, Sousa OE, Schofield CJ, Dujardin JP, Guhl F, Mas-Coma S 2008. Phylogeography and genetic variation of *Triatoma dimidiata*, the main Chagas disease vector in Central America, and its position within the genus *Triatoma*. *PLoS Negl Trop Dis* 2: e233.
- Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW 2012. GenBank. *Nucleic Acids Res* 40: D48-53.
- Dorn PL, Calderon C, Melgar S, Moguel B, Solorzano E, Dumonteil E, Rodas A, de la Rúa N, Garnica R, Monroy C 2009. Two distinct *Triatoma dimidiata* (Latreille, 1811) taxa are found in sympatry in Guatemala and Mexico. *PLoS Negl Trop Dis* 3: e393.
- Dorn PL, Monroy C, Curtis A 2007. *Triatoma dimidiata* (Latreille, 1811): a review of its diversity across its geographic range and the relationship among populations. *Infect Genet Evol* 7: 343-352.
- Fanello C, Santolamazza F, della Torre A 2002. Simultaneous identification of species and molecular forms of the *Anopheles gambiae* complex by PCR-RFLP. *Med Vet Entomol* 4: 461-464.
- Herrera-Aguilar M, Be-Barragan LA, Ramirez-Sierra MJ, Tripet F, Dorn P, Dumonteil E 2009. Identification of a large hybrid zone between sympatric sibling species of *Triatoma dimidiata* in the Yucatan peninsula, Mexico, and its epidemiological importance. *Infect Genet Evol* 9: 1345-1351.
- Marcella A, Bargues MD, Ramsey JM, Magallon-Gastelum E, Salazar-Schettino PM, Abad-Franch F, Dujardin JP, Schofield CJ, Mas-Coma S 2001. The ITS-2 of the nuclear rDNA as a molecular marker for populations, species and phylogenetic relationships in Triatominae (Hemiptera: Reduviidae), vectors of Chagas disease. *Mol Phylogenet Evol* 18: 136-142.
- Mas-Coma S, Bargues MD 2009. Populations, hybrids and the systematic concepts of species and subspecies in Chagas disease triatomine vectors inferred from nuclear ribosomal and mitochondrial DNA. *Acta Trop* 110: 112-136.
- Thompson JD, Higgins DG, Gibson TJ 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
- Van Bortel W, Trung HD, Roelants P, Harbach RE, Backeljau T, Coosemans M 2000. Molecular identification of *Anopheles minimus* s.l. beyond distinguishing the members of the species complex. *Insect Mol Biol* 9: 335-340.