

## RESEARCH

## Phylogenetics Based on Partial ORF2 of Triatoma Virus in Triatomines Collected Over a Decade From Domiciliary Habitats

María Laura Susevich,<sup>1</sup> Gerardo Aníbal Marti,<sup>1</sup> Agustín Balsalobre,<sup>1</sup> and María Gabriela Echeverría<sup>2,3</sup><sup>1</sup>Centro de Estudios Parasitológicos y de Vectores (CEPAVE-CCT-La Plata-CONICET-UNLP), 2 #584, (1900) La Plata, Argentina<sup>2</sup>Cátedra de Virología, Faculty of Veterinary Sciences (UNLP), 60 y 118 (1900) La Plata, Argentina<sup>3</sup>Corresponding author, e-mail: mariagabrielaecheverria@yahoo.com.ar

Subject Editor: Seth Barribeau

J. Insect Sci. 15(17): 2015; DOI: 10.1093/jisesa/ieu139

**ABSTRACT.** The only virus sequenced and studied in triatomines is the Triatoma virus, from the Dicistroviridae family, which causes delayed development, reduced oviposition, and premature death of infected insects. With the goal of expanding the sequences already obtained in previous years and verifying if any changes occurred in their genomic sequences, 68 samples of triatomines from several provinces of Argentina were analyzed. Sixteen positive samples were obtained by Reverse Transcription (RT)-polymerase chain reaction using the VP3-VP1 subregion of open reading frame-2 as a diagnostic method; after sequencing, 11 samples were obtained from *Triatoma infestans*. These new sequences showed no significant differences in the analyzed regions, which were not grouped by species or habitat or geographical distribution. There were no differences when compared with the sequences found during 2002–2012, all obtained from the wild. We conclude that despite being an RNA virus, the different sequences show high homology.

**Key Words:** *Triatoma infestans*, domiciliary habitat, triatoma virus, Argentina

To date, Triatoma virus (TrV) is the only entomopathogenic virus found and identified in triatomines (Muscio et al. 1987). It is a member of the Dicistroviridae family, whose type species is the cricket paralysis virus. TrV is a single-stranded RNA virus consisting of 9,010 bases, which replicates in intestinal epithelial cells causing delayed development and death of infected insects (Muscio et al. 1997). The TrV genome has two open reading frames (ORF1 and ORF2, with 5,387 and 2,606 nucleotides [nt], respectively). These ORFs are separated by an intergenic region of 172 nt. ORF1 is located between nt 549 and 5,936 and codes for nonstructural proteins, whereas ORF2 is located between nt 6,109 and 8,715 (Czibener et al. 2000). ORF2 codes for the structural proteins of the viral capsid are VP2, VP4, VP3, and VP1. Because of its high pathogenicity and vertical transmission, TrV is considered a potential agent for biological control of its host *Triatoma infestans* (Muscio et al. 1997), the vector of the protozoan parasite *Trypanosoma cruzi* that causes Chagas disease in humans. About 8–11 million people are estimated to be infected with this disease in Latin America. Consequently, TrV has the potential to be exploited to control its disease-bearing hosts (Gordon and Waterhouse 2006).

In 1987, the incidence of TrV in populations of *T. infestans* in homes and around houses in different localities of the endemic area in Córdoba province was studied, showing an infection rate of 1–10% (Muscio et al. 1997). During a survey in the 2002–2005 period, TrV was again found in natural populations of *T. infestans* in Córdoba province and also in the provinces of San Luis, Mendoza, Tucumán, La Rioja, Santa Fe, Santiago del Estero, and Salta. In Chaco province, *Triatoma sordida* was found and cited as the second species observed in nature (Marti et al. 2009). In 2012, three species from natural areas of Chaco province, *T. infestans*, *T. delpontei*, and *Psammolestes coreodes*, were found to be positive for TrV (Susevich et al. 2012). At present, TrV has only been found in Argentina in nature.

The goal of this study was to expand the search for TrV in triatomines from domiciles and peridomiciles from different provinces of Argentina and to study virus diversity according to geographic distribution and species by means of the analysis of genomic sequences.

## Materials and Methods

***T. infestans* Samples Collection.** Samples of *T. infestans* were collected from different regions in 10 provinces of Argentina (Buenos Aires, Chaco, Córdoba, Formosa, La Rioja, Mendoza, San Luis, Santa Fe, Santiago del Estero, and Tucumán) between 2002 and 2012. Searches were conducted in peridomiciles, cracks in adobe walls as well as in every deposit, corrals, etc. Dislodging substances such as tetramethrin 0.2% were used to facilitate collection of the insects when necessary. The insects collected individually were transported to the laboratory in sterile plastic containers, each containing a folded piece of paper and capped with fine mesh screen. Subsequently, the insects were identified following Lent and Wygodzinsky (1979) and maintained at a temperature of  $28 \pm 1^\circ\text{C}$ , with  $60 \pm 5\%$  relative humidity and a photoperiod of 12:12 (L:D) h in special Centro de Estudios Parasitológicos y de Vectores (CEPAVE) facilities. There were a total of 68 samples that were stored at  $-80^\circ\text{C}$  until use. Each sample contained 10 triatomines fecal material from each region.

**Processing of Samples and RNA Extraction.** Approximately 10 mg of fecal sample from the triatomines was homogenized with 200  $\mu\text{l}$  of phosphate-buffered saline (PBS) in microtubes. RNA was extracted directly from fecal samples. A volume of 50  $\mu\text{l}$  of samples resuspended in PBS was homogenized in 300  $\mu\text{l}$  of TRIZOL reagent (Gibco-Invitrogen, M & M section), according to the manufacturer's instructions. RNA concentration was determined by measuring absorbance at 260 nm (A260) in a spectrophotometer (Marti et al. 2008).

**RT-Polymerase Chain Reaction (PCR) as Diagnostic Method.** This first RT-PCR, called "diagnostic-PCR," used primers TrV1a (5' TCA AAATAACTATCATTCTGG 3') and TrV1b (5' TTCAGCCT TATCCCCCCC 3'), covering a region between VP3 and VP1, with an expected size of 832 bp. RT-PCR was performed following the methodology of Marti et al. (2008). The first TrV strain obtained from *T. infestans* captured in peridomiciliary environments in 2001 (Guanaco Muerto Dean Funes, Cordoba Province) was used as positive control (Susevich et al. 2012).

**Purification and Sequence Analysis.** PCR products were purified using Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). Both chains of each product were sequenced using the same primers as for PCR using the BigDye Terminator V3.1 Cycle Sequencing Kit (Applied Biosystems) in an ABI3130XL Sequencer (Applied Biosystems), Unidad Genómica, INTA Castelar, Argentina. The sequences were edited using the software BioEdit version 5 (Hall 1999), and homology analyses were performed with the BLASTN program (<http://www.ncbi.nlm.nih.gov/BLAST/>). The sequences were then aligned by means of the MEGA program version 4.0 using the ClustalW algorithm. The nucleotide sequences for each subregion were translated to their corresponding amino acids with DNAMAN version 4.13. The sequences were then aligned and compared with the reference sequence and other sequences (Fig. 1). The sequence pair distances were calculated by DNASTAR (Tamura et al. 2007). The phylogenetic trees were constructed using the MEGA program with the neighbor joining (NJ) method, and bootstrap analyses were conducted using 1,000 replicates. Evolutionary distances were computed using the maximum composite likelihood method (Tamura et al. 2007). To perform phylogenetic reconstruction of amino acids, 15 taxa were used with the NJ method and Bootstrap of 1,000 replicates with Poisson correction.

**Sequences Included in This Work.** Some previously published sequences were included to verify differences in host, environment, and geographical distribution. Names, province, year, and accession numbers are AF178440 (Czibener et al. 2000), TIN1-ARG (Córdoba 2002), TIN2-ARG (Chaco 2007), TDE1-ARG (Chaco 2007) and PCO1-ARG (Chaco 2005)—GenBank HM044313, HM044314, HM044312, and HM044315, respectively—(Susevich et al. 2012) and one sequence of *T. delpontei* isolated from a bird nest (Psittaciformes) from Chaco province in 2007 (TDE2-ARG accession number JQ657727).

## Results

**Diagnostic PCR.** In total, 68 samples were analyzed. The “diagnostic-PCR” using primers TrV1a and TrV1b gave the expected size of 832 bp. Sixteen samples were positive by this RT-PCR, and 11 of these samples could be sequenced—TIN3 to TIN13—(Table 1).

**Sequence Analysis.** The sequences obtained were aligned and compared with each other and to the reference strain (AF178440) as well as to the sequences obtained in previous studies (Susevich et al. 2012). No deletions or insertions were detected, only point mutations were observed. The percentage of sequence identity ranged from 99.3 to 100% (10 nucleotide changes in positions 86, 222, 223, 344, 374, 449, 484, 494, 572, and 750). The highest level of nucleotide identity was found in three groups: between PCO1-ARG and TIN2-ARG; between TDE2-ARG and TIN3-ARG; and one large group formed by TDE1-ARG, TIN5-ARG, TIN6-ARG, TIN8-ARG, TIN11-ARG, and TIN13-ARG. The amino acid sequence identity ranged from 98.9 to 100% (four amino acid changes in positions 74, 161, 191, and 250) (Table 2). A 100% level of amino acid identity was found in two groups: one

consisting of AF178440, PCO1-ARG, TIN2-ARG, and TIN10-ARG and the other formed by TDE1-ARG, TDE2-ARG, TIN3-ARG, TIN5-ARG, TIN6-ARG, TIN7-ARG, TIN8-ARG, TIN11-ARG, and TIN13-ARG.

**Phylogenetic Analysis.** The region analyzed for phylogenetic reconstruction comprised TrV1a and TrV1b. All the sequences obtained were deposited in GenBank as follows: TIN3-ARG (JQ657728), TIN4-ARG (KC618414), TIN5-ARG (JX963130), TIN6-ARG (KC618411), TIN7-ARG (JX963131), TIN8-ARG (JX963132), TIN9-ARG (JX963133), TIN10-ARG (JX963134), TIN11-ARG (JX963135), TIN12-ARG (KC618412), and TIN13-ARG (KC618413) (Table 1). As shown in Fig. 1, two discrete clusters were formed with high bootstrap values and divided in two branches each: one included *T. infestans* and other triatomine species (*P. coreodes* PCO1) and the other included *T. infestans* and *Triatoma delpontei* (TDE1 and TDE2). Overall, the sequences did not tend to segregate according to triatomine species, year, or geographical origin.

## Discussion

Our results indicate the widespread distribution of the virus, not only in wild species (Susevich et al. 2012) but also in peridomestic environment, thus increasing the likelihood of its potential use as a biological control agent against the vectors of Chagas disease. To date, the programs to control American trypanosomiasis are based exclusively on the use of chemicals in and around households, and no test has been conducted using TrV in the wild. The elucidation of their diversity and number, hosts, and pathogenic effects may reveal major roles of dicistroviruses in the biosphere (Bonning and Miller 2010). The prevalence of TrV in Argentina has been partially studied; however, in other areas, endemic for Chagas disease, such as other Latin America countries, is still unknown.

Structural and nonstructural viral proteins have been shown to be suitable targets for phylogenetic studies in other dicistroviruses. This study presents a phylogenetic analysis of the structural protein gene region of TrV sequences. We focused our investigations on the structural protein gene because this genomic region usually shows a higher degree of divergence than nonstructural genomic regions. Comparison of the complete genome sequences of central European acute bee paralysis virus (ABPV) genotypes and the reference strain did not result in significant divergences at any regions of the investigated strains (Bakonyi et al. 2002). Other investigations have shown that the Kashmir bee virus and ABPV of the honeybee can be broadly separated by different continents of origin but that it is more difficult to identify regional trends within each continent. Although nonstructural proteins

**Table 1. Sequenced samples of *T. infestans* with their location, year, and the accession number of the nucleotide sequence**

Name	Province/collection year	Accession number
TIN3-ARG	Santa Fe/2009	JQ657728
TIN4-ARG	Tucumán/2003	KC618414
TIN5-ARG	Mendoza/2003	JX963130
TIN6-ARG	La Rioja/2004	KC618411
TIN7-ARG	Santa Fe/2010	JX963131
TIN8-ARG	Tucumán/2003	JX963132
TIN9-ARG	San Luis/2003	JX963133
TIN10-ARG	Tucumán/2003	JX963134
TIN11-ARG	La Rioja/2004	JX963135
TIN12-ARG	S. del Estero/2003	KC618412
TIN13-ARG	La Rioja/2003	KC618413

**Table 2. Amino acid changes in the subregion of ORF2 from the sequences of Argentine TrV, with respect to reference strain AF178440 (3)**

Subregion	Position	AA	Sequence
TrV 1a/1b	74	R	TIN9
		K	AF178440, TIN2, TIN4, TIN10, PCO1
		E	TIN1, TIN3, TIN5, TIN6, TIN7, TIN8, TIN11, TIN12, TIN13, TDE1, TDE2
	161	A	AF178440, TIN1, TIN2, TIN3, TIN4, TIN5, TIN6, TIN7, TIN8, TIN9, TIN10, TIN11, TIN13, PCO1, TDE1, TDE2
		V	TIN12
		F	TIN4
	191	L	AF178440, TIN1, TIN2, TIN3, TIN5, TIN6, TIN7, TIN8, TIN9, TIN10, TIN11, TIN12, TIN13, PCO1, TDE1, TDE2
		V	TIN1
		L	AF178440, TIN2, TIN3, TIN4, TIN5, TIN6, TIN7, TIN8, TIN9, TIN10, TIN11, TIN12, TIN13, PCO1, TDE1, TDE2, TIN12

can be expected to have less variability within the genome, this is not always the rule (De Miranda et al. 2004). In the case of the black queen cell virus (BQCV), one study showed that the 5'-proximal third of ORF1 was the most variable region and contains several insertions and deletions (Reddy et al. 2013). However, other study showed that ORF1 or ORF2 should be useful to classify BQCV genotypes according to geographical origin (Noh et al. 2013). On the other hand, ORF2 showed better phylogenetic grouping corresponding to the geographical origin of the genotypes as in the case of ABPV (Tapaszti et al. 2009). These authors performed a phylogenetic analysis based on the structural proteins of ORF2 of BQCV (514 bp), which provided higher resolution than ORF1 and were able to separate strains from Poland, Austria, and Hungary. Moreover, other authors found two distinct lineages by analyzing the intergenic region and part of the ORF2 of Israeli Acute Paralysis Virus (Blanchard et al. 2008). At present, we are not able to compare sequences from different origins, because TrV has not yet been described in countries other than Argentina in nature. There is insufficient sequence data available from the database.

With respect to the sequence analysis described in this study, mutations were found at nucleotide level resulting in amino acid changes at various positions. In the region covered by the primers and TrV1a and TrV1b, changes consisted of 10 nt, corresponding to four amino acid changes. All the sequences obtained in this study were closely related, even though they had been isolated from different triatomines, different environments and provinces, and in different years.

We have been faced with the difficulty of the low rate of recovery of RNA from the stool samples of the triatomines analyzed. We attribute this primarily to the fact that, despite the extreme care taken to manipulate the materials to minimize the risk of contamination, which causes RNA to be inactivated, the latter could not be recovered in its entirety from the samples preserved in nuclease free water at  $-80^{\circ}\text{C}$ . Stool samples, although extremely useful, are not available for subsequent uses. Moreover, these observations raise the question of the specificity of the primers used or whether the RNA obtained from different samples was not obtained as single strand.

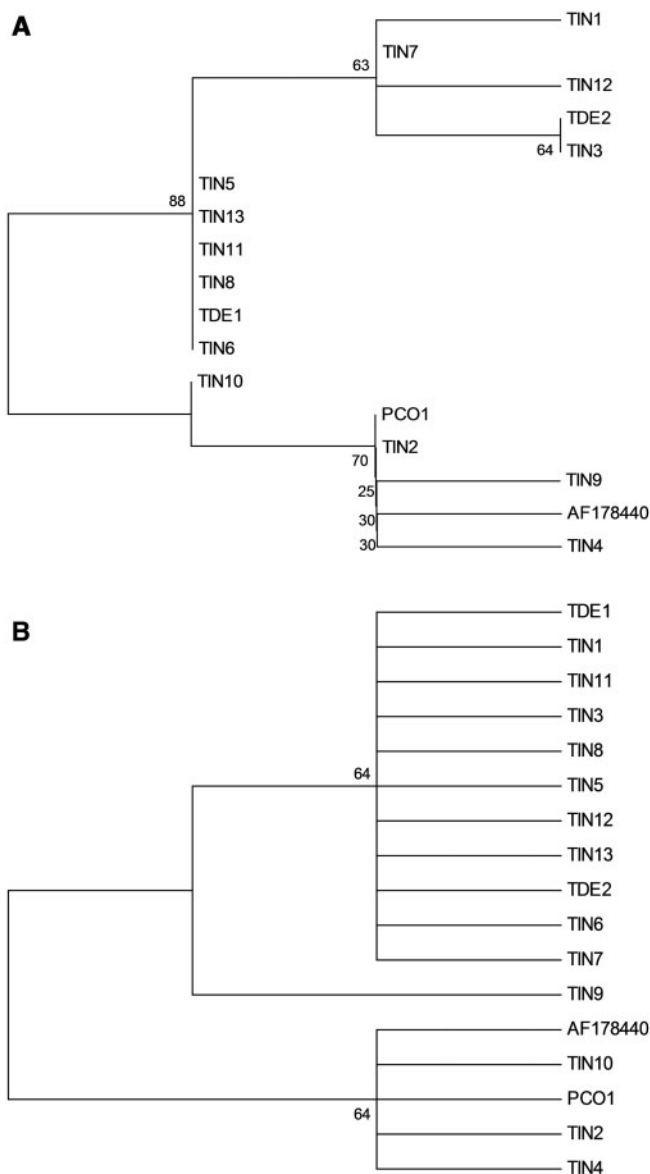
Because there is no clear geographical, temporal, or ecological separation between the sequences to support their clear phylogenetic separation based on ORF2 in our limited number of sequences obtained, future research into RT-PCR-based diagnostics should target regions of the ORF1 or the intergenic region, because these may have a high level of variation and allow us to classify genotypes according to geographical origin. A more comprehensive genome analysis should be performed in the future to identify and classify TrV isolates from different geographic regions. Moreover, further studies are required to increase our knowledge and understanding about the presence of TrV in Latin American countries.

#### Acknowledgments

This study was partially supported by the following grants: PIP 0007 and PICT No. 2011-1081 ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica), Argentina, and CYTED (Programa Iberoamericano de Ciencia y Tecnología para el desarrollo) (www.cytcd.org) 209RT0364 AE-2009-1-21. We thank the Cátedra de Virología, Facultad de Ciencias Veterinarias, Universidad Nacional de La Plata, Argentina, and the CEPAVE (Centro de Estudios Parasitológicos y de Vectores) and CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) where this work was developed. We also thank Unidad de Epidemiología Molecular, Instituto de Patología Experimental, Facultad de Ciencias de la Salud, Universidad Nacional de Salta, CONICET (Salta, Argentina), and to the Centro de Referencia de Vectores, Coordinación Nacional de Control de Vectores (Santa María de Punilla, Córdoba, Argentina).

#### References Cited

- Bakonyi, T., E. Grabensteiner, J. Kolodziejek, M. Rusvai, G. Topolska, W. Ritter, and N. Nowotny. 2002. Phylogenetic analysis of acute bee paralysis virus strains. *Appl. Environ. Microbiol.* 68: 6446–6450.
- Blanchard, P., F. Schurr, O. Celle, N. Cougoule, P. Drajnudel, R. Thiery, J. Faucon, and M. Ribiere. 2008. First detection of Israeli acute paralysis virus (IAPV) in France, a dicistrovirus affecting honeybees (*Apis mellifera*). *J. Invertebr. Pathol.* 99: 348–350.
- Bonning, B. C., and A. W. Miller. 2010. Dicistroviruses. *Annu. Rev. Entomol.* 55: 129–150.
- Czibener, C., J. L. La Torre, O. A. Muscio, R. A. Ugalde, and E. A. Scodeller. 2000. Nucleotide sequence analysis of *Triatoma* virus shows that it is a member of a novel group of insect RNA viruses. *J. Gen. Virol.* 81: 149–1154.
- De Miranda, J. R., M. Drebot, S. Tyler, M. Shen, C. E. Cameron, D. B. Stoltz, and S. M. Camazine. 2004. Complete nucleotide sequence of Kashmir bee virus and comparison with acute bee paralysis virus. *J. Gen. Virol.* 85: 2263–2270.
- Gordon, K.H.J., and P. M. Waterhouse. 2006. Small RNA viruses of insects: expression in plants and RNA silencing, pp. 459–502. *In* B. C. Bonning (ed.), *Insect viruses: biotechnological applications*. Elsevier, San Diego, CA.



**Fig. 1.** Phylogenetic trees obtained by the NJ method from the analysis of nucleotides (A) and putative amino acids (B) from the TrV1a and TrV1b subregion of Argentine and reference sequences. Numbers above branches indicate support values.

- Hall, T. A. 1999.** BioEdit: A user friendly biological sequences alignment editor and analysis program for Window 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95–98.
- Lent, H., and P. Wygodzinsky. 1979.** Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas diseases. *Bull. Am. Mus. Nat. Hist.* 163: 123–520.
- Marti, G., E. T. González, J. J. Garcia, A. R. Viguera, D.M.A. Guerin, and M. G. Echeverría. 2008.** AC-ELISA and RT-PCR assays for the diagnosis of *Triatoma virus* (TrV) in triatomines (Hemiptera: Reduviidae) species. *Arch. Virol.* 153: 1427–1432.
- Marti, G. A., M. G. Echeverría, M. L. Susevich, J. J. Becnel, S. A. Pelizza, and J. J. García. 2009.** Prevalence and distribution of parasites and pathogens of Triatominae from Argentina, with emphasis on *Triatoma infestans* and *Triatoma virus* TrV. *J. Invertebr. Pathol.* 102: 233–237.
- Muscio, O. A., J. L. La Torre, and E. A. Scodeller. 1987.** Small nonoccluded viruses from triatomine bug *Triatoma infestans* (Hemiptera: Reduviidae). *J. Invertebr. Pathol.* 49: 218–220.
- Muscio, O. A., J. L. La Torre, M. A. Bonder, and E. A. Scodeller. 1997.** *Triatoma virus* pathogenicity in laboratory colonies of *Triatoma infestans* (Hemiptera: Reduviidae). *J. Med. Entomol.* 34: 253–256.
- Noh, J. H., K. E. Reddy, S. E. Choe, M. S. Yoo, H.T.T. Doan, C. H. Kweon, M. Ramya, B. S. Yoon, L.T.K. Nguyen, T.T.D. Nguyen, et al. 2013.** Phylogenetic analysis of black queen cell virus genotypes in South Korea. *Virus Genes* 46: 362–368.
- Reddy, K. E., J. H. Noh, S. E. Choe, C. H. Kweon, M. S. Yoo, H.T.T. Doan, M. Ramya, B. S. Yoon, L.T.K. Nguyen, T.T.D. Nguyen, et al. 2013.** Analysis of the complete genome sequence and capsid region of black queen cell viruses from infected honeybees (*Apis mellifera*) in Korea. *Virus Genes* 47: 126–132.
- Susevich, M. L., G. A. Marti, M. S. Serena, and M. G. Echeverría. 2012.** New *Triatoma virus* hosts in wild habitats of Argentina. *J. Invertebr. Pathol.* 110: 405–407.
- Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007.** MEGA4: molecular evolutionary genetic analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596–1599.
- Tapaszti, Z., P. Forgách, C. Kovago, G. Topolska, N. Nowotny, M. Rusvai, and T. Bakonyi. 2009.** Genetic analysis and phylogenetic comparison of Black queen cell virus genotypes. *Vet. Microbiol.* 139: 227–234.

Received 29 October 2013; accepted 29 January 2014.