Infection, Genetics and Evolution 43 (2016) 274–280

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

The nuclear elongation factor-1α gene: a promising marker for phylogenetic studies of Triatominae (Hemiptera: Reduviidae)

Sebastián Díaz, Omar Triana-Chávez, Andrés Gómez-Palacio *

Grupo de Biología y Control de Enfermedades Infecciosas - BCEI, Universidad de Antioquia (UdeA), Medellín, Antioquia, Colombia

ARTICLE INFO

Article history:
Received 24 February 2016
Received in revised form 31 May 2016
Accepted 3 June 2016
Available online 4 June 2016

Keywords:
Triatominae
Chagas disease
Phylogeny
EF-1α

ABSTRACT

Molecular systematics is a remarkable approach for understanding the taxonomic traits and allows the exploration of the inter-population dynamics of several species in the Triatominae subfamily that are involved in *Trypanosoma cruzi* transmission. Compared to other relevant species that transmit vector-borne diseases, such as some species of the Diptera, there are relatively few nuclear genetic markers available for systematic studies in the Triatominae subfamily. Molecular systematic studies performed on Triatominae are based on mitochondrial gene fragments and, less frequently, on nuclear ribosomal genes or spacers. Due to the fact that these markers can occasionally present problems such as nuclear mitochondrial genes (NUMTs) or intra-genomic variation for high gene copy numbers, it is necessary to use additional nuclear markers to more reliably address the molecular evolution of Triatominae. In this study, we performed phylogenetic analysis using the nuclear elongation factor-1 *alpha* (*EF-1α*) gene in individuals from 12 species belonging to the Triatomini and Rhodniini tribes. Genetic diversities and phylogenetic topologies were compared with those obtained for the mitochondrial 16S rRNA and Cytochrome b (*cyt b*) genes, as well as for the D2 variable region of the ribosomal 28S rRNA gene. These results indicate that the *EF-1α* marker exhibits an intermediate level of diversity compared to mitochondrial and nuclear ribosomal genes, and that phylogenetic analysis based on *EF-1α* is highly informative for resolving deep phylogenetic relationships in Triatominae, such as tribe or genera.

1. Introduction

The Triatominae subfamily (Hemiptera: Reduviidae) is a taxonomic group containing insects that are involved in transmitting the parasite *Trypanosoma cruzi* to humans and other mammals. The subfamily comprises around 146 species, 15 genera and 5 tribes (Abad-Franch et al., 2013; Ayala, 2009; Da Rosa et al., 2012; Frías-Lasserre, 2010; Gonçalves et al., 2013; Jurberg et al., 2009; Schofield and Galvão, 2009). Most of these are distributed throughout the Neotropical and subtropical zones, with the exception of some *Triatoma* species and the *Linshcosteus* genus, which are also present in East Asia, where the epidemiological relevance of these insects has recently been increasing (Dujardin et al., 2015).

Molecular systematic studies proposed for Triatominae have been based on gene trees that use fast evolving molecular markers, such as the mitochondrial genes 12S rRNA, 16S rRNA, *COI* and *cyt b* (Justi et al., 2014; Lyman et al., 1999; Monteiro et al., 2004); the nuclear ribosomal genes 18S rRNA and 28S rRNA; and ITS regions (Bargues et al., 2008; Hypsa et al., 2002; Justi et al., 2014). Although there is a generally accepted consensus about the usefulness of these molecular markers, it is well-known that in arthropods, mitochondrial and nuclear ribosomal markers can present problems that lead to misinterpretation, due to the presence of nuclear mitochondrial genes (NUMTs) (Rogers and Griffiths-Jones, 2012) or intra-genomic variation for high copy numbers of ribosomal clusters (Bargues et al., 2014; Bower et al., 2009). In this sense, systematic studies of triatomines based exclusively on mitochondrial and rRNA genes are not always appropriate for all systematic levels (i.e. discussed in Mas-Coma and Bargues, 2009).

The elongation factor-1 *alpha* (*EF-1α*) gene has been well-characterized in other arthropods and has been used for several systematic studies in several taxonomic groups (Djernæs and Damgaard, 2006). Unlike rRNA genes, *EF-1α* is a single-copy coding region that presents a moderate diversity level due to its exon-intron structure (Djernæs and Damgaard, 2006). Although paralogous copies have been described in species from the Hemiptera: Coccoidea superfamily (Downie and Gullan, 2004), these can be identified according to exon-intron structure of the gene (Djernæs and Damgaard, 2006); moreover, there is no evidence of paralogous copies in other groups within Hemiptera to date. In this paper, we assess the use of the *EF-1α* gene for phylogenetic studies in the Triatominae subfamily by analyzing several species from

* Corresponding author at: Grupo de Biología y Control de Enfermedades Infecciosas - BCEI, Universidad de Antioquia, Calle 70 # 52–21, Medellín, Colombia.
E-mail address: amgomezpa@gmail.com (A. Gómez-Palacio).

http://dx.doi.org/10.1016/j.meegid.2016.06.010
1567-1348/

the tribes Triatomini and Rhodniini and comparing the phylogenetic signal with some of the most frequently used molecular markers to date.

2. Materials and methods

2.1. Insects, DNA extraction and PCR-amplification of EF-1α

A total of 31 individuals from nine species were processed for *EF-1α* amplification (Table 1). Genomic DNA was extracted from the legs or thoraxes of individual insects, as described elsewhere in the literature (Collins et al., 1987). From each specimen, we amplified a 737 bp fragment of the *EF-1α* gene using the primers TEF1-F (5′-CAGGCCGATTGTGCKGTACTKAT-3′) and TEF2-R (5′-GCTTCATGRTG CATTTCAAC-3′). These primers were adapted from those previously designed for Gerridae (Damgaard and Sperling, 2001), by including the publicly available *EF-1α* transcripts for *R. prolixus* (Genbank code: JO495068) and *T. matogrossensis* (Genbank code: HP429357). PCR reactions were conducted in a final volume of 35 μl, using 2 μl of 50-ng/μl DNA, 1X PCR buffer (0.1 M Tris–HCl, 0.5 M KCl and 0.015 M $MgCl_2$, pH 8.3), 0.2 mM dNTP, 0.5 μM of each primer, 3 mM $MgCl_2$ and 1.25 U/μl of Taq DNA polymerase (Fermentas®). Amplifications consisted of an initial denaturation cycle at 95 °C for 5 min, and 35 cycles at 95 °C for 60 s, 54 °C for 35 s, and 72 °C for 60 s, followed by a final extension cycle at 72 °C for 10 min. In order to compare *EF-1α* with other mitochondrial and nuclear markers, in some species we also amplified a 682 bp fragment of the *cyt b* gene, as well as 629 bp of the D2 variable region of the ribosomal gene 28S rRNA (D2-28S rRNA). *Cyt b* gene was PCR-amplified using primers CYTB7432 (5′-GGACGWGGWA TTTATTATGGATC-3′) and CYTB7433 (5′-GCWCCAATTCARGTTARTAA-3′) (Lyman et al., 1999). PCR reactions were conducted in a final volume of 35 μl using 30-ng of DNA templates, 1X PCR buffer (0.1 M Tris–HCl, 0.5 M KCl, and 0.015 M MgCl2, pH 8.3), 250-μM dNTP, 0.016-μM of each primer, 35-mM MgCl2 and 2 U of Taq DNA polymerase. The fragments were amplified with the following thermal cycling conditions: 95 °C for 5 min; 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 60 s; 72 °C for 10 min. D2-28S rRNA was PCR-amplified using the primers D2F (5′–GCGAGTCGTGTTGCTTGATAGTGCAG-3′) and D2R (5′–TTGGTCCGTGTTTCAAGACGGG − 3′) as described elsewhere (Porter and Collins, 1996). PCR reactions were conducted in a final volume of 35 μl using 30-ng of DNA templates, 1X PCR buffer (0.1 M Tris–HCl, 0.5 M KCl, and 0.015 M $MgCl_2$, pH 8.3), 250-mM dNTP, 0.025-mM of each primer, 3-mM $MgCl_2$ and 2 U of Taq DNA polymerase (Fermentas®). After an initial denaturation of 95 °C for 5 min, PCR reactions were subjected to 35 cycles at 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, followed by a final extension cycle at 72 °C for 7 min (Herrera-Aguilar et al., 2009).

All PCR products were sequenced by Macrogen Inc. (http://www.macrogen.com). Forward and reverse sequences were trimmed according to a Q-value >20 (Q-score is analogous to Phred score) to obtain high quality sequences (Q-value = −10log 10 (PE), where PE is the probability of error and Q = 20 corresponds to 99% accuracy). Both sequences were aligned to obtain a consensus sequence for each individual, using the default parameters in ClustalX (Thompson et al., 1997), as implemented in BioEdit v.7.0.9.0 (Hall, 1999). All sequences used in this study are available in Genbank (Table 1).

2.2. Sequence analyses

Publicly available *EF-1α* nucleotide sequences for *T. brasiliensis*, *T. matogrossensis* and *R. prolixus* were included in the analyses, and the complete dataset was compared against the mitochondrial 16S rRNA and *cyt b* genes, as well as against the nuclear D2-28S rRNA gene (Table 1).

For each dataset, multiple alignments were performed using the auto strategy, as implemented in MAFFT v.7 (Katoh and Standley, 2013). For the *EF-1α* fragment used, the exon-intron structure was defined by comparing it with the *Drosophila melanogaster EF-1α* gene (Genbank code: X06869.1) and the genome location was identified in the *Rhodnius prolixus* genome (Genbank assembly code: GCA_000181055.3). The rate of recombination per gene R (Hudson, 1987) and the minimum number of recombination events RM (Hudson and Kaplan, 1985) were estimated for the amplified *EF-1α* fragment using DnaSP v.5.10.01 (Librado and Rozas, 2009). The percentage of nucleotide variability (S/L: number of variable sites over alignment size), nucleotide composition (%GC), overall mean standard K2P genetic distance, and overall frequency for transitional/transversional substitutions (R = ts/tv) were calculated using MEGA v.6.0.6 (Tamura et al., 2013).

2.3. Phylogenetic analyses

In order to assess the phylogenetic suitability of the markers used, the level of substitution saturation was tested by analyzing the pairwise ts/tv versus inter-species pairwise K2P divergence. The global index for the observed substitution saturation (Iss) (Xia et al., 2003) was estimated using DAMBE v.5.2.63 (Xia, 2013). The dataset for each marker was used to infer the best-fitting model with the Bayesian information criterion (BIC) (Schwarz, 1978), as implemented in MEGA v.6.0.6 (Tamura et al., 2013). Phylogenies for *EF-1α* were inferred with intron regions both included and excluded. Maximum likelihood (ML) phylogenies were inferred using the Tamura-Nei (TN93 + G + I) substitution model for *cyt b*, the Hasegawa-Kishino-Yano (HKY + G) model for 16S rRNA, the Tamura 3-parameter (T92 + G) for D2-28S rRNA, and (T92 + I) for EF-1α, respectively. In all cases, the Nearest-Neighbor-Interchange

Table 1
Triatominae species studied.

Tribe	Genera	Species (n)	Code in Fig. 1	Marker			
				EF-1α	D2-28S rRNA	16S rRNA	*cyt b*
Triatomini	*Triatoma*	*T. maculata* (13)	*Tmac*	KX109892	KX109904	AY035465	KX109901
		T. brasiliensis	*Tbra*	EC915954	GQ853395	AF021183	AY336524
		T. infestans	*Tinf*	KX109893	GQ853397	AF021198	AF045721
		T. matogrossensis	*Tmat*	HP429357	GQ853398	AY035454	KC249272
		T. dimidiata (6)	*Tdim*	KX109894	KX109905	AY035448	AF045726
	Panstrongylus	*P. lignarius*	*Plig*	KX109895	KX109906	JQ897823	KX109902
		P. geniculatus	*Pgen*	KX109896	KX109907	AF394593	KX109903
Rhodniini	*Rhodnius*	*R. pallescens* (6)	*Rpal*	KX109897	KC543522	AF045706	EU363823
		R. colombiensis	*Rcol*	KX109898	KC543516	AY035438	FJ229360
		R. ecuadoriensis	*Recu*	KX109899	KC543517	AF045711	AF045715
		R. prolixus	*Rpro*	JO495068	AF435862	AF324519	AF045718
		R. barretti	*Rbar*	KX109900	–	–	JX273159
Ougroup	*Reduvius*	*R. personatus*	*Rper*	DQ531735	–	AY318879	–
	Stenopoda	*S. cinerea*	*Scin*	–	FJ230730	–	GQ869681