Experimental Parasitology 122 (2009) 328–337

Contents lists available at ScienceDirect

Experimental Parasitology

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

Trypanosoma cruzi: New insights on ecophylogeny and hybridization by multigene sequencing of three nuclear and one maxicircle genes

M. Subileau [a], C. Barnabé [b], E.J.P. Douzery [c], P. Diosque [d], M. Tibayrenc [b,*]

[a] *UMR 1083 "Sciences pour l'Oenologie", équipe "Microbiologie et Technologie des Fermentations", ENSAM/INRA, 2 place Viala, 34060 Montpellier Cedex 01, France*
[b] *UR 165 Génétique et Evolution des Maladies Infectieuses (GEMI), UMR CNRS/IRD 2724, Institut de Recherche pour le Développement, 911 Av. Agropolis, BP 64501, 34394 Montpellier Cedex 5, France*
[c] *Laboratoire de Paléontonlogie, Paléobiologie et Phylogénie CC064, Institut des Sciences de l'Evolution de Montpellier (UMR 5554 CNRS), Université Montpellier II, Montpellier, France*
[d] *Unidad de Epidemiología Molecular (UEM), Instituto de Patología Experimental, Universidad Nacional de Salta, Av. Bolivia 5150, CP4400, Salta, Argentina*

ARTICLE INFO

Article history:
Received 3 January 2009
Received in revised form 7 April 2009
Accepted 12 April 2009
Available online 21 April 2009

Keywords:
Trypanosoma cruzi
Kinetoplastidae
Multigene sequencing
Phylogeny
Hybrid origins
Recombination
Clonality
Discrete typing units

ABSTRACT

Natural populations of *Trypanosoma cruzi* are structured into five genetic lineages, *T. cruzi* I and *T. cruzi* II a to e, as the result of clonal evolution with rare genetic recombination events. To explore more in depth these phenomenons, a multigene sequencing approach was used, for the first time in the case of *T. cruzi*. Three nuclear loci and a maxicircle locus were sequenced on 18 *T. cruzi* stocks. Sequences were used to build phylogenetic trees from each locus and from concatenated sequences of all loci.

The data confirmed the hybrid origin of DTUs IId and IIe, as the result of an ancient genetic recombination between strains pertaining to IIb and IIc. The data confirmed also a hybrid origin of DTUs IIa and IIc. Contrary to previous reports, we failed to detect mosaic genes. The phylogenetic relationship between DTUs and the respective roles of recombination and selection were tested.

Note: Nucleotide sequence data reported in this paper are available in the GenBank™ database under the accession numbers EU186432–EU186535.

1. Introduction

Chagas disease remains a serious health problem in Latin America where it is the main cause of cardiopathy. Over 18 million people are afflicted in these endemic regions with 13,000 casualties reported each year (WHO, 2003). *Trypanosoma cruzi* is transmitted to humans and other mammals by the feces of a hematophagous triatomine bug, by blood transfusion and congenitally from infected mothers (Prata, 2001). Two transmission cycles of *T. cruzi* interact: a sylvatic one involving different species of triatomine bugs and wild mammals, and a domestic one in which humans and domestic mammals act as reservoirs. The distinction between these two cycles is based on geographical distribution of triatomines and vertebrate hosts, and is associated to the preferences of insects for specific blood sources.

Trypanosoma cruzi natural populations are very heterogeneous genetically and phenotypically. They have a predominant clonal structure, with only occasional bouts of genetic exchange, (de Lana et al., 1998; Cruz et al., 2006). Analyses based on multilocus enzyme electrophoresis (MLEE) (Barnabé et al., 2000), random amplified polymorphic DNA (RAPD) (Brisse et al., 2000a), and structural and functional variability of small subunit rRNA and mini exon genes (Souto et al., 1996; Brisse et al., 2001) have shown that natural clones of *T. cruzi* are distributed into two main lineages called *T. cruzi* I and *T. cruzi* II (or TC I and TC II). TCII is itself subdivided into lesser lineages IIa, b, c d and e. These lineages cannot be equated to clades, since it is now admitted that hybridization and recombination play a notable role in *T. cruzi* long-term molecular evolution. However these subdivisions appear to be very reliable and stable in space and time. The term "discrete typing unit" (DTU) has been coined to describe such sets of pathogen strains that are genetically closer to each other than to any other stock, are stable in space and time, and are identifiable by common genetic, molecular, or immunological markers called «tags» (Tibayrenc, 1998), *T. cruzi* DTUs show a strong statistical tendency to exhibit distinct biological properties (Laurent et al., 1997; Revollo et al., 1998) and show variations in virulence (Montamat et al., 1996; Brenière et al., 1998). *T. cruzi* I is widespread on the whole

* Corresponding author. Fax: +33 467416299.
E-mail address: Michel.Tibayrenc@ird.fr (M. Tibayrenc).

0014-4894/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2009.04.008

endemic zone, from The United States to Argentina, and is found in both sylvatic and domestic cycles. *T. cruzi* II predominates in southern cone countries of America, although it has been found in North American sylvatic cycles too (Barnabé et al., 2001). DTUs IIa and IIc are mainly encountered in the sylvatic cycle, although IIa has been found in Ecuatorian domestic cycles (Garzon et al., 2002). DTUs IIb, IId and IIe are mostly domestic with IId and IIe being particularly associated to the vector species *Triatoma infestans* and to humans (Barnabé et al., 2000). Hybrid genotypes were first evidenced in this parasite's natural populations with MLEE and RFLP profiles (Bogliolo et al., 1996), and with RAPD markers (Carrasco et al., 1996). DTUs IId and IIe were proved to be products of an ancient hybridization event between DTUs IIb and IIc (Barnabé et al., 2000, 2003). Westenberger et al. (2005) have postulated that genotypes IIa and IIc on one hand, IId and IIe on the other hand, are the results of hybridization events between TC I and IIb, and IIb and IIc, respectively.

Experimental genetic exchanges were performed with genetically modified strains of *T. cruzi* I, leading to the production of hybrids, similar to natural hybrids (Gaunt et al., 2003). These experimental results suggest that, at least *T. cruzi* I, has an extent capacity for genetic exchanges under various processes, including fusion of parental genotypes, loss of alleles, rare homologous recombination and uniparental inheritance of kinetoplastic maxicircle DNA.

This study was undertaken to apply a Multigene sequencing approach to the agent of Chagas disease. Gene sequencing is highly resolutive since each bp carries a bit of information, and is highly portable, since sequences can be exchanged between laboratories through email. However the sequence of unique genes is hardly representative of the overall phylogeny of the whole species (Nichols, 2001). On the other hand, multilocus typing approaches such as MLEE or RAPD do not convey this drawback, but are more subject to homoplasy. Multigene sequencing combines the advantages of both approaches.

Differently from the MLST (Multilocus Sequence Typing) widely applied to bacteria (Maiden et al., 1998), our study has a purpose of basic research (phylogeny, mode of reproduction), rather than a molecular epidemiology goal. In the present study, we selected three nuclear gene fragments (Leucine Aminopeptidase [*Lap*], Superoxide dismutase a and b [*SodA* and *SodB*]) and one kinetoplastic gene fragment (the subunit 1 of NADH dehydrogenase [*Nd1*]), according to previous MLEE results (Tibayrenc et al., 1986; Barnabé et al, 2000). In MLEE experiments, *Lap* exhibits a high polymorphism and an apparent lack of heterozygous genotypes, while *SodA* and *SodB* show a high power to discriminate between TC I from TC II). It should be however underlined that MLEE phenotypes do not have a 100% predictive value about the sequences of the genes that code for them. *Nd1* was selected to have a nonnuclear gene that makes it possible to analyse a different mode of evolutionary pattern, and to compare our result to previous one (Brisse et al., 2003). This gene sample is limited by comparison with the MLST approach, which typically relies on the sequencing of fragments of 6–10 housekeeping genes (Maiden et al., 1998). However, again, our goal is basic evolutionary analysis rather than molecular epidemiology. In this perspective, this study based on the sequencing of four genes already is quite informative.

Eighteen laboratory-cloned *T. cruzi* stocks, three from each DTUs, were sequenced. Evolution of the species was surveyed through three different aspects: phylogenetic trees, recombination and analysis of natural selection. Indeed, while recombination has been investigated by many authors, selection, which can play an important evolutionary role, especially in the event of positive selection, has been much less explored.

2. Materials and methods

2.1. Parasites

Origins and previous characterizations of the stocks are given in Table 1. All stocks have been laboratory-cloned, under microscopic control. Genotypes and DTU attribution of the stocks have been previously determined by MLEE (Barnabé et al., 2000), RAPD (Brisse et al., 2000a) and/or sequence-characterized amplified region (SCAR) markers (Brisse et al., 2000b). Parasite culture and DNA preparation were as described (Barnabé et al., 2000).

2.2. Multilocus sequencing

Four molecular markers were isolated. For the leucine aminopeptidase gene fragment, the 486-nt fragment of nuclear *Lap* gene (GenBank AI066263, (Porcel et al., 2000)) was used to design primers (primer L: 5′-TGT ACA TTG TGC TTG GCT GAG-3′ and primer R: 5′-GCT GAG GTG ATT AGC GAC AAA-3′). For superoxide dismutase A, the 649-nt fragment of *SodA* (GenBank AI623018) was used (L: 5′-CCA CAA GGC GTA TGT GGA C-3′ and R: 5′-GTG GAA CGG CAA AGC AGT-3′). For superoxide dismutase B, the 894-nt fragment of *SodB* (GenBank U90723, (Ismail et al., 1997)) was used (L: 5′-GCC CCA

Table 1
Geographic, host origins and previous characterization of the 18 *Trypanosomes cruzi* stocks and 2 *T. cruzi marinkellei* stocks under study.

Stock	Species DTU	Country	Locality	Host
P209 cl93	*T. cruzi* I	Bolivia	Sucre	Human
CUTIA cl1	*T. cruzi* I	Brazil	Colatina	*Dasyprocta aguti*
OPS21 cl11	*T. cruzi* I	Venezuela	Macuayas	Human
CAN III cl1	*T. cruzi* IIa	Brazil	Belem	Human
Dog Theis	*T. cruzi* IIa	USA	Oklahoma	*Canis familiaris*
92122102R	*T. cruzi* IIa	USA	Statesboro	*Procyon lotor*
MAS1 cl1	*T. cruzi* IIb	Brazil	Brasilia	Human
IVV cl4	*T. cruzi* IIb	Chile	Cuncumen	Human
TU18 cl93	*T. cruzi* IIb	Bolivia	Tupiza	*Triatoma infestans*
M5631 cl5	*T. cruzi* IIc	Brazil	Island of Marajo	*Dasypus novemcinctus*
M6241 cl6	*T. cruzi* IIc	Brazil	Belem	Human
X109/2	*T. cruzi* IIc	Paraguay	Makthlawaiya	*Canis familiaris*
MN cl2	*T. cruzi* IId	Chile	Illapel	Human
SC43 cl1	*T. cruzi* IId	Bolivia	Santa Cruz	*Triatoma infestans*
Bug 2148 cl1	*T. cruzi* IId	Brazil	Rio Grande do Sul	*Triatoma infestans*
CL Brener	*T. cruzi* IIe	Brazil	Rio Grande do sul	*Triatoma infestans*
P63 cl1	*T. cruzi* IIe	Paraguay	Makthlawaiya	*Triatoma infestans*
Tula cl2	*T. cruzi* IIe	Chile	Tulahuen	Human
B3	*T. c. marinkellei*	Brazil	Sao Felipe	*Phyllostomus discolor*
B7 cl11	*T. c. marinkellei*	Brazil	Sao Felipe	*Phyllostomus discolor*