Available online at www.sciencedirect.com
SCIENCE DIRECT
Experimental Parasitology 113 (2006) 1–7
Experimental Parasitology
www.elsevier.com/locate/yexpr

Phylogenetic analysis of the glucose-6-phosphate isomerase gene in *Trypanosoma cruzi* ☆

Hélène Broutin *, Frédérique Tarrieu, Michel Tibayrenc, Bruno Oury [1], Christian Barnabé

Institut de Recherche pour le Développement (IRD), Génétique des Maladies Infectieuses, Unité de Recherche IRD no 165/Unité Mixte de Recherche CNRS (Centre National de la Recherche Scientifique)-IRD no 2724, 911 Avenue Agropolis, BP 64501, 34394 Montpellier cedex 5, France

Received 13 September 2005; received in revised form 17 November 2005; accepted 21 November 2005
Available online 10 January 2006

Abstract

Trypanosoma cruzi, the agent of Chagas disease, has a basically clonal population structure with rare hybridization events. The species is subdivided into six "Discrete Typing Units" called DTUs I, IIa–e, distributed into two major phylogenetic lineages, *T. cruzi* I and II (TC I and II). The glucose-6-phosphate isomerase (Gpi) is a specific isoenzymic locus that presents homozygous profiles for DTUs I, IIa–c, and typical heterozygous patterns, for DTUs IId and IIe. The gene was sequenced in 12 *T. cruzi* stocks and in three stocks pertaining to related species. The phylogenetic relationships observed confirm that the DTUs I, IIa–c do constitute monophyletic groups. Nevertheless, the phylogenetic hierarchy of the DTUs is not clearly resolved with the GPI gene. The hybrid status of DTUs IId and IIe was clearly supported. Sequence analysis revealed that the allele *4* present in both DTUs IIa and IIc, previously considered as unique, displayed in fact two distinct sequences, specific for each DTU. The level of recombination between alleles has been investigated.

Index Descriptors and Abbreviations: *Trypanosoma cruzi*; Glucose-6-phosphate isomerase locus; Discrete typing unit; Hybrid; Intraspecific phylogeny; Recombination

1. Introduction

Trypanosoma (*Schizotrypanum*) *cruzi* (Chagas, 1909) the protozoan parasite responsible for Chagas disease in South and Central America, shows considerable genetic polymorphism. Analysis of multilocus enzyme electrophoresis (MLEE) polymorphism in population genetics terms has revealed that *T. cruzi* undergoes predominant clonal evolution (Tibayrenc et al., 1986). This model does not rule out the possibility of occasional bouts of sexuality, and states only that such events are too rare to break the prevalent pattern of clonal population structure (Tibayrenc and Ayala, 1988). Moreover, it has been proposed that *T. cruzi* clonal genotypes are distributed into two major genetic subdivisions (Souto et al., 1996; Tibayrenc, 1995) which have been referred to as *T. cruzi* I (TC I) and *T. cruzi* II (TC II) (Momen, 1999). However, occurrence of recombination events has been confirmed both in the field (Bogliolo et al., 1996; Brisse et al., 1998; Carrasco et al., 1996; Machado and Ayala, 2001), and in the laboratory (Gaunt et al., 2003). This makes it impossible to equate these genetic subdivisions to real clades. This is a general methodological difficulty in many pathogenic microorganisms that undergo both clonal evolution and horizontal gene transfer (Tibayrenc, 1998). To solve this problem, the concept of "Discrete Typing Unit" (DTU) was proposed (Tibayrenc, 1998) to designate discrete sets of genotypes that are genetically more related to each other than to any other genotypes, and that can be characterized by specific molecular markers called "tags." Studies based on MLEE and RAPD markers have allowed the

☆ Nucleotide sequence data reported in the paper are available in the GenBank databases under the Accession Nos. AY484472–AY484486.

* Corresponding author. Fax: +04 67 41 62 99.
E-mail address: broutin@mpl.ird.fr (H. Broutin).
[1] Present address: Institut de Recherche pour le Développement (IRD), Unité de Recherche no 8 "Pathogénie des *Trypanosomatidae*," BP 64501, 34394 Montpellier cédex 5, France.

0014-4894/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2005.11.014

identification of six DTUs, of which DTU I corresponds to *T. cruzi* I, and DTUs IIa–e are lesser subdivisions of *T. cruzi* II (Barnabé et al., 2000; Brisse et al., 2000a).

Glucose-6-phosphate isomerase (EC 5.3.1.9, GPI), is an important enzyme involved in glycolysis that catalyzes the reversible isomerization of D-glucose 6-phosphate to D-fructose 6-phosphate. The different genotypes of the corresponding isoenzyme locus, *Gpi*, are DTU-specific (tags). The observed genotypes correspond to homozygous profiles for DTU I (*5/5*), IIa and IIc (*4/4*), and IIb (*3/3*), and typical heterozygous patterns for DTUs IId and IIe (*2/4* and *3/4*, respectively) (Barnabé et al., 2000).

A phylogenetic analysis of MLEE and RAPD variability permitted the proposal that DTUs IId and IIe correspond to hybrid genotypes whose putative parents would be genotypes pertaining to DTUs IIb and IIc (Barnabé et al., 2000; Brisse et al., 2000b, 2003; Machado and Ayala, 2001). These hybrid genotypes appear to be stabilized by subsequent clonal propagation. They seem to have a high epidemiological and medical relevance, since they have been isolated from many human cases of Chagas disease, and are distributed over wide geographical areas ("major clones" (Tibayrenc and Ayala, 1988)). To explore more in depth this hypothesis of hybridization, we have sequenced and compared the different *Gpi* alleles in a set of stocks representative of the six *T. cruzi* DTUs, and in two related species (*T. cruzi marinkellei* and *Trypanasoma rangeli*) taken as comparison.

2. Materials and methods

2.1. Parasite stocks

Origins of the stocks are given in Table 1. The 12 *T. cruzi* stocks are representative of the whole phylogenetic diversity of the species (two stocks for each DTU) and three stocks from two related species were taken as potential outgroups. All the stocks have previously been laboratory-cloned, with visual control of the presence of a single cell and previously characterized by MLEE studies (Barnabé et al., 2000). Parasite cultures and DNA preparations were as described (Brisse et al., 2000a).

2.2. Amplification of the Gpi gene and sequencing

Primers were selected within homologous regions of previously reported sequences: the *Gpi* gene sequences of *T. cruzi* (strain CL Brener) (GenBank entry: AF052833) (Andersson et al., 1998) and *T. brucei* (GenBank entry: X15540) (Marchand et al., 1989) were aligned. Primers were selected within homologous regions, namely gpi1 starting at the position 370 (5′-CTTCATATTGCTTTGCGAA ATCG-3′) and gpi2 starting at the position 1126 (5′-TTCC CCAGTTCCACACCCCA-3′). A 1274 bp fragment representing 70% of the gene length was amplified in each stock. The PCRs were performed in a final volume of 100 µl containing 10 mM Tris–HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3, 4× 5 mM dNTP, 2× 500 nM primers, and 1 U *Taq* DNA polymerase (Roche Applied Science, Mannheim, Germany) and 200 ng genomic DNA. Thirty-five cycles (denaturation: 1 min at 94 °C; annealing: 1 min at 59 °C; elongation: 1 min 30 at 72 °C) were followed by a final elongation step of 5 min at 72 °C. The amplified products were detected after electrophoresis on 1.5% agarose gel in TAE 0.5× buffer (40 mM Tris–acetate, 2 mM EDTA, pH 8.3) and staining with ethidium bromide. The fragments were purified with the Kit "Wizzard PCR Preps DNA purification system" (Promega). Double strand sequencing of PCR products was carried out by using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Sequences showing ambiguous base calls, mostly between two bases indicating possible heterozygote, were cloned in pGEM-T Easy Vector System (Promega) according to the manufacturer's instructions.

2.3. Phylogenetic analysis

Sequences were edited and manually aligned with Seaview (Galtier et al., 1996). The maximum likelihood method (ML) (Felsenstein, 1981) was used because of its advantage:

Table 1
Geographic and host origins of the 15 *Trypanosoma* stocks under study

Strain	Species (DTU)	Country/locality	Host
OPS21cl11	*T. cruzi* (I)	Venezuela/Cojedes, Macuayas	Human
P209cl1	*T. cruzi* (I)	Bolivia/Sucre	Human
CANIIIcl1	*T. cruzi* (IIa)	Brazil/Belém	Human
Ep-272	*T. cruzi* (IIa)	Colombia/Porvenir, Meta	*Rhodnius prolixus*
CBBcl3	*T. cruzi* (IIb)	Chile/Coquimbo, Tulahuén	Human
Tu18cl92	*T. cruzi* (IIb)	Bolivia/Tupiza	*Triatoma infestans*
M6241cl6	*T. cruzi* (IIc)	Brazil/Belém	Human
X110/8	*T. cruzi* (IIc)	Paraguay/Makthlawaiya	Dog
Mncl2	*T. cruzi* (IId)	Chile/Coquimbo	Human
Bug2148cl1	*T. cruzi* (IId)	Brazil/Rio Grande do Sul	*Triatoma infestans*
CL Brener	*T. cruzi* (IIe)	Brazil/Rio Grande do Sul	*Triatoma infestans*
Tulacl2	*T. cruzi* (IIe)	Chile/Coquimbo, Tulahuén	Human
B3	*T. c. marinkellei*	Brazil/Bahia, Sao Felipe	*Phyllostomum discolor*
B7	*T. c. marinkellei*	Brazil/Bahia, Sao Felipe	*Phyllostomum discolor*
RGB	*T. rangeli*	Venezuela, Caracas	Dog

by using an explicit model of nucleotide substitution, it accounts for transition/transversion bias, unequal nucleotide frequencies and among-site rate variation (Yang, 1996). The optimal model of nucleotide substitution was determined by using Modeltest v.3.06 (Posada and Crandall, 1998), which tests the goodness of fit of 14 models (with or without invariants and/or Gamma distribution) with the likelihood ratio test (Felsenstein, 1988). All trees were reconstructed by using PAUP (Swofford, 2003), version 4β 10. Robustness of nodes in the trees was evaluated by bootstrap (Felsenstein, 1985) computing 100 resamplings by heuristic ML reconstruction. The presence of internal recombination events (mosaic genes) was tested by the RDP v.2 program (Martin et al., 2005), which attempts to identify recombination breakpoints using 10 published recombination detection methods, including Geneconv, Bootscan, Maximun chi2, Chimaera, and Sister Scanning.

To test whether positive selection is operating on the gene, we compared the relative abundance of synonymous and nonsynonymous substitutions within the sequences. For a pair of sequences, this is done by first estimating the number of synonymous substitutions per synonymous site (dS) and the number of nonsynonymous substitutions per nonsynonymous site (dN), and their variances: Var(dS) and Var(dN), respectively. With this information, we can test the null hypothesis that H0: $d\mathrm{N} = d\mathrm{S}$ using a Z test, where $Z = (d\mathrm{N} - d\mathrm{S})/\mathrm{Sqrt}(\mathrm{Var}(d\mathrm{S}) + \mathrm{Var}(d\mathrm{N}))$. This test, as well as translation of DNA sequences into polypeptides, was performed by using MEGA version 3.1 (Kumar et al., 2004).

3. Results

3.1. Sequence polymorphism analysis

PCR products from the 15 stocks under study could be double-strand sequenced from 1192 to 1235 nucleotides and manually aligned without indel over 1192 nucleotides. Seven sequences showed nucleotide ambiguities between two bases, at 27 different loci (see Table 2). The corresponding PCR products were cloned in a plasmid vector and at least one allele was sequenced, the sequence of the other allele was deduced in most cases. The sequence of *T. rangeli* RGB was removed from the study because it was too distant from the other sequences and probably not correctly aligned (mean of identity with the other sequences was 0.85, compared to 0.96 with *T. cruzi marinkellei* and to 0.98, which was the mean of identity within *T. cruzi*). Out of 1192 characters, 1103 were constant, 89 were variable, in which 11 were variable and parsimony-uninformative and 78 were variable and parsimony-informative. The frequencies of bases were 0.287, 0.192, 0.235, and 0.286 for A, C, G, and T, respectively, show-

Table 2
Sequences at the 27 sites showing uncertainties between two bases for the 14 stocks under study

Group	*T. cruzi* I		*T. cruzi* IIa		*T. cruzi* IIb		*T. cruzi* IIc		*T. cruzi* IId		*T. cruzi* IIe		*T. c. marinkellei*	
Stock sites:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
7	A	A	G	G	A	A	G	G	G	G	A	A/G	A	A
28	C	C	C	C	T	T	C	A	A/T	A/T	A/T	A/T	T	T
61	C	C	C	C	C	C	C	A	A/C	A/C	A/C	A/C	C	C
87	A	A	A	A	A	A	G	G	A/G	A/G	A/G	A/G	A	A
108	A	A	A	A	A	A	G	G	A/G	A/G	A/G	A/G	A	A
177	T	T	T	T	T	T	T	C	C/T	C/T	C/T	C/T	T	T
182	A	A	A	A	G	G	A	A	A/G	A/G	A/G	A/G	A	A
183	G	G	G	G	G	G	A	A	A/G	A/G	A/G	A/G	G	G
234	A	A	A	A	A	A	A	A	A	A	A/G	A/G	A	A
258	A	A	A	A	T	T	A	A	A/T	A/T	A/T	A/T	A	A
324	T	T	T	T	C	C	T	T	C/T	C/T	C/T	C/T	T	T
330	T	T	T	T	C	C	T	T	C/T	C/T	C/T	C/T	C	C
347	G	G	G	G	A	A	G	G	A/G	A/G	A/G	A/G	A	A
348	T	T	T	T	T	T	T	T	G/T	G/T	G/T	G/T	T	T
349	G	G	G	G	G	G	G	G	A/G	A/G	A/G	A/G	G	G
350	A	A	A	A	A	A	A	A	A/G	A/G	A/G	A/G	C	C
351	G	G	G	G	G	G	G	G	G/T	G/T	G	G/T	G	G
375	G	G	G	G	G	G	A	A	A/G	A/G	A/G	G	G	G
378	C	C	C	C	T	T	C	C	C/T	C/T	T	C/T	C	C
417	A	A	G	G	G	G	A	A	A/G	G	A/G	A/G	A	A
464	C	C	C	C	A	A	C	C	A/C	A/C	A/C	A/C	C	C
570	C	C	T	T	G	G	C	C	C/G	C/G	C/G	C/G	A/G	A/G
693	C	C	T	T	T	T	C	C	C/T	C/T	C/T	C/T	T	T
741	A	A	A	A	A	A	A	G	A/G	A/G	A/G	A/G	A	A
829	G	G	G	G	G	G	G	G/T	G	G	G	G	G	G
1080	C	T	C	C	G	G	C	C	G	G	C/G	C/G	C	C
1146	C	C	C	C	T	T	C	C	C/T	C/T	C/T	C/T	C	C

Stocks: 1, OPS21cl11; 2, P209cl1; 3, CANIIIcl1; 4, Ep-272; 5, CBBcl3; 6, TU18; 7, M6241; 8, X110.8; 9, Mncl2; 10, Bug2148cl1; 11, CLBrener; 12, Tulacl2; 13, *T.c.m.*B7; 14, *T.c.m.*B3.

ing a low C+G content (0.427) for this gene. These frequencies are equally distributed among the stocks.

3.2. Phylogenetic analysis

The best-fit model of substitution for this gene, selected by Modeltest v. 3.06, was the Hasegawa–Kishino–Yano's model (HKY), using variable base frequencies and variable transition and transversion frequencies, with gamma distributed rate variation among sites (G), and proportion of invariable sites (I). This model (HKY+I+G) was used to perform a heuristic search of the maximum likelihood trees by using a *T. cruzi marinkellei* stock as outgroup and only one tree was found (Fig. 1). Four groups of stocks were supported by high level of bootstrap, DTU I (bootstrap value (BP) = 100), DTU IIa (BP = 100), DTU IIb with alleles from DTUs IId-IIe (BP = 95) and DTU IIc with the other alleles from DTU IId-IIe (BP = 99). The cluster 1 (Fig. 1, top) was not supported by a high bootstrap value (75) nor was the cluster DTU I-IIa (55). The heuristic search of the maximum parsimony trees gave similar results (data not shown): three trees of equal length (122 steps) with the same four principal groups supported by significant bootstrap values (>95%). DTU IIb (with one of the alleles of hybrid stocks) was in external position in all trees whereas the relative positions of the three other groups varied among the trees.

3.3. Uncertainties between two bases

Ninety-six ambiguous base calls, only between two bases, were found at 27 sites (see Table 2), with an unequal repartition: 8 A or C (M), 43 A or G (R), 8 A or T (W), 6 C or G (S), 23 C or T (Y), and 8 G or T (K). Ninety-three of them (97%) were found in the two hybrid groups DTUs IId and TC IIe; 1 was found in X110.8 stock, 1 in both B3 and B7 *T. c. marinkellei* stocks. In 22 sites over the 26 containing two bases in hybrid stocks, one base was found in DTU IIb and the other one was found in DTU IIc; in the other four sites with uncertainties, only one of the two bases was encountered in both DTUs IIb and IIc. In DTU IIc, X110/8 stock seemed to be the most probable parent whereas in DTU IIb, both stocks, showing the same sequence at these sites, could equally be one of the parents. It is worth noting that all the uncertainties have been resolved by molecular cloning in plasmid vectors.

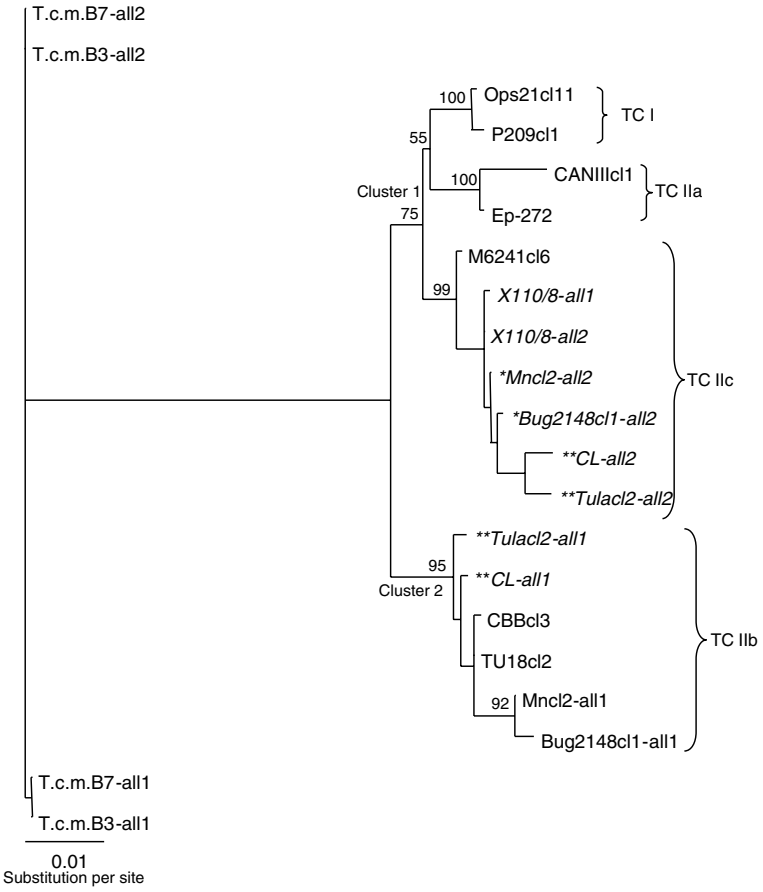

Fig. 1. Maximum likelihood phylogram of 21 alleles from 12 *Trypanosoma cruzi* stocks and 2 *T. c. marinkellei* stocks, reconstructed from the Gpi partial gene ($\ln L = 2406.9$). The model of nucleotide substitution model used was the HKY85+G+I model. Bootstrap percentages obtained by ML after 100 replicates are given for each node. This is an unrooted tree rooted by outgroup method (here the 2 *T.c.m.* stocks was used as outgroup). Note that for the 4 *T. cruzi* stocks from hybrid DTUs, where two alleles were detected in the sequences, one allele, arbitrarily named "all1" falls into TC IIb group whereas the second one, named "all2," falls into the TC IIc group.

3.4. Recombination detection

Over the six methods automatically used by the program RDP2, only two of them (Siscan and MaxChi) detected recombination breakpoints. In all cases, only one parent could be identified with certainty, the other parent being a putative one. The putative daughter sequences (mosaic) identified were Ops21cl11 (with parent (p) Bug-all1, and putative parent (pp) Bug-all2); P209cl1 (p X110/8-all1, pp Bug-all1); CANIIIcl1 (p Bug-all1, pp P209cl1); M6241cl6 (p P209cl1, pp CANIIIcl1); Mncl2-all2 (p Bug-all1, pp Ops21cl11); CL-all2 (p TU18cl2, pp CANIIIcl1); and Tulacl2-all2 (p P209cl1, pp CANIIIcl1).

For all these potential recombination events, the program indicated possible misidentifications of beginning breakpoints, and it is worth noting that the other four systems (namely Rdp, Geneconv, Bootscan, and Chimaera) failed to detect any recombination event. Finally, when the program listed only the events detected by more than two methods, no recombination event was found.

3.5. Synonymous and non-synonymous mutations and protein polymorphism analysis

The overall mean distance based on the number of mutations (calculated pairwise, with all *T. cruzi* and *T. c. marinkellei* sequences) was 18.95, and 9.77 for synonymous and nonsynonymous mutations, respectively. Within *T. cruzi* only (i.e., without *T. c. m. marinkellei* sequences), these means were 9.74 and 8.27 for synonymous and nonsynonymous mutations, respectively. The Z test of selection, performed on this last set of sequences, rejected H0, which is $d\mathrm{N} = d\mathrm{S}$, significantly ($p < 0.001$) and indicated a possible purifying selection, where $d\mathrm{N} < d\mathrm{S}$, for this gene. Out of the 397 amino acids obtained by translation, 371 were constant, 26 were variable, in which 22 were parsimony-informative and 4 parsimony-uninformative. Pairwise distances based on number of amino acid substitutions varied from 0 to 15, with an overall average of 8.81. A neighbour joining tree based on these distances corroborated the ML tree based on nucleotide sequences (data not shown).

4. Discussion

In all trees, already known genetic distances between species were corroborated (i.e., *T. rangeli* very far from *T. cruzi* and *T. cruzi marinkellei* close to *T. cruzi* but always separated from it, in external position) (Barnabé et al., 2003). Nevertheless, sequences of *T. rangeli* seemed too divergent to consider the alignment as correct. Despite the likely conservative constraints for a protein coding region and a possible purifying selection for this gene that eliminates deleterious mutations from a population, the *Gpi* gene remains highly variable within *T. cruzi* species and appears to be a reliable marker to explore the phylogenetic relationships within the species *T. cruzi*, including bat trypanosomes, but not much further. Within *T. cruzi*, both the maximum likelihood tree (ML-tree) and the maximum parsimony trees (MP-tree) confirmed that DTUs I, IIa, IIb, and IIc are robust monophyletic units, as evidenced by highly significant bootstrap values. This corroborates results from other authors based on the study of two nuclear single-copy genes and one maxicircle sequence (Machado and Ayala, 2001). The external position of DTU IIb was previously observed with mitochondrial genes (Machado and Ayala, 2001), with a divergence estimated to about 10 Myr assuming a 1% substitution rate per million years (Brisse et al., 2003). Such a result was not observed in various studies, using MLEE, RAPD, rDNA or miniexon markers (Barnabé et al., 2000; Brisse et al., 2000a; Souto et al., 1996), which have been used by a college of specialists to identify two main groups in *T. cruzi*, *T. cruzi* I (TC I), and *T. cruzi* II (TC II). These discrepancies could not be solely explained by evolutionary differences between nuclear DNA vs. mitochondrial DNA, because an external position of DTU IIb sequences has been previously found in experimental hybrids with different nuclear genes including *Gpi* (Gaunt et al., 2003). Obviously, recombination exists and was reported (Gaunt et al., 2003; Machado and Ayala, 2001; Sturm et al., 2003), each author proposing an evolutionary scenario to explain incongruities between trees from different markers. A consensus tends to be clear now that not only DTUs IId and IIe are hybrid (see below) but also that DTUs IIa/IIc (formerly called Miles'Z3, (Miles et al., 1978)) could have an ancient hybrid origin, by recombination between TC I and II (Sturm et al., 2003).

The analysis of recombination of the *Gpi* gene, reported here, did not confirm these scenarios. Stocks from DTUs IIa/IIc (CANIIIcl1 and M6241cl6) showed some putative mosaic genes with DTUs I and IIb as putative parents, but these mosaics are probably false because only two methods showed recombination points over six methods scanned by the powerful program RDP (i.e., SiScan and MaxChi), while Rdp, Bootscan, Geneconv, and Chimaera failed to detect recombination breakpoints (Martin et al., 2005). Moreover, when we aligned the *Gpi* mosaic sequence (GenBank Accession No. AY227811, (Gaunt et al., 2003)), the program RDP failed to detect any recombination event in the 942 bp common part of the sequence. Nevertheless, recombination could be located in other parts of the sequence not used by us, or/and the putative parents be absent in our sample.

All the alleles from DTUs IId and IIe clearly clustered with either DTU IIb or IIc. This is the clear confirmation of the previous hypothesis of hybrid origin of DTUs IId and IIe, between stocks from DTUs IIb and IIc (Barnabé et al., 2000; Brisse et al., 2000a) now widely confirmed by other authors (Gaunt et al., 2003; Machado and Ayala, 2001; Sturm et al., 2003). Interestingly, none of these alleles clustering with IIb or IIc, was very distant to homozygous alleles of pure IIb/IIc genotypes, which is confirmed by significant bootstrap values in both groups. Like for IIa/IIc pure genotypes, alleles from heterozygote clustering with IIc or IIb showed a low level of recombination, detected by few methods, with parents belonging

to DTUs I, IIb, and interestingly IIa, that could suggest a new type of recombination IIa/IIc (see Section 3 for details). These data tend to confirm the unusual mode of recombination in *T. cruzi* probably consisting in a fusion of parental genotypes, homologous recombination, but with a low frequency for the *Gpi* gene, and probably loss of alleles in the course of the time, as previously demonstrated (Gaunt et al., 2003). It is worth noting that the stocks belonging to the hybrid genotypes DTUs IId/IIe are frequently isolated from chagasic patients (Barnabé et al., 2000), which makes the study of their evolutionary mode medically relevant in MLEE studies, Gpi bands from DTUs IIa and IIc migrate at the same electrophoretic distance and have been consequently interpreted as a unique allele. Here, we demonstrated that the gene sequences from DTUs IIa and IIc, were in fact distinct. This isoenzyme band corresponds actually to two different proteins with very similar or identical global electric charges that cannot be separated by electrophoresis under our technical conditions. Complete sequencing is more resolutive than isoenzyme analysis, which is well known. Allozymes should be considered as families of closely related alleles rather than as real alleles.

The results reported here do not basically question the actual knowledge about the genetic structure of *T. cruzi* species; nevertheless they bring new insight about levels of recombination between alleles of an important housekeeping gene, the *Gpi*, of which the protein isoenzymic pattern is often used for rapid identification of *T. cruzi* strains, in routine in many laboratories. The six DTUs previously defined by our team and others (Barnabé et al., 2000; Brisse et al., 2000a,b) remain stable and reliable entities in spite of recombination events regularly reported. Their stability, including those genotypes that are supposed to have a hybrid origin, has been verified over more than 20 years (Tibayrenc et al., 1981). They can therefore be considered as robust units of analysis for all applied studies dealing with *T. cruzi*, including experimental evolution, epidemiological tracking, vaccine and drug design. In the same way, the classification into two main groups (TC I and TC 2) is not questioned in spite of the external position of DTU IIb for certain genes including *Gpi*.

Acknowledgment

Authors were supported by the Institut de Recherche pour le Développement (IRD).

References

Andersson, B., Aslund, L., Tammi, M., Tran, A.N., Hoheisel, J.D., Pettersson, U., 1998. Complete sequence of a 93.4-kb contig from chromosome 3 of *Trypanosoma cruzi* containing a strand-switch region. Genome Research 8, 809–816.

Barnabé, C., Brisse, S., Tibayrenc, M., 2000. Population structure and genetic typing of *Trypanosoma cruzi*, the agent of Chagas disease: a multilocus enzyme electrophoresis approach. Parasitology 120, 513–526.

Barnabé, C., Brisse, S., Tibayrenc, M., 2003. Phylogenetic diversity of bat trypanosomes of subgenus *Schizotrypanum* based on multilocus enzyme electrophoresis, random amplified polymorphic DNA, and cytochrome *b* nucleotide sequence analyses. Infection Genetics and Evolution 2, 201–208.

Bogliolo, A.R., Lauriapires, L., Gibson, W.C., 1996. Polymorphisms in *Trypanosoma cruzi*: evidence of genetic recombination. Acta Tropica 61, 31–40.

Brisse, S., Barnabe, C., Banuls, A.L., Sidibe, I., Noel, S., Tibayrenc, M., 1998. A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. Molecular and Biochemical Parasitology 92, 253–263.

Brisse, S., Barnabe, C., Tibayrenc, M., 2000a. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. International Journal of Parasitology 30, 35–44.

Brisse, S., Dujardin, J.C., Tibayrenc, M., 2000b. Identification of six *Trypanosoma cruzi* lineages by sequence-characterised amplified region markers. Molecular And Biochemical Parasitology 111, 95–105.

Brisse, S., Henriksson, J., Barnabe, C., Douzery, E.J., Berkvens, D., Serrano, M., De Carvalho, M.R., Buck, G.A., Dujardin, J.C., Tibayrenc, M., 2003. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. Infection Genetics and Evolution 2, 173–183.

Carrasco, H.J., Frame, I.A., Valente, S.A., Miles, M.A., 1996. Genetic exchange as a possible source of genomic diversity in sylvatic populations of *Trypanosoma cruzi*. American Journal of Tropical Medicine and Hygiene 54, 418–424.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17, 368–376.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 1988. Phylogenies from molecular sequences: inference and reliability. Annual Review of Genetics 22, 521–565.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Computing Applied Bioscience 12, 543–548.

Gaunt, M.W., Yeo, M., Frame, I.A., Stothard, J.R., Carrasco, H.J., Taylor, M.C., Mena, S.S., Veazey, P., Miles, G.A., Acosta, N., De Arias, A.R., Miles, M.A., 2003. Mechanism of genetic exchange in American trypanosomes. Nature 421, 936–939.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinformatics 5, 150–163.

Machado, C.A., Ayala, F.J., 2001. Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. Proceedings of the National Academy of Sciences of the United States of America 98, 7396–7401.

Marchand, M., Kooystra, U., Wierenga, R.K., Lambeir, A.M., Van Beeumen, J., Opperdoes, F.R., Michels, P.A., 1989. Glucosephosphate isomerase from *Trypanosoma brucei*. Cloning and characterization of the gene and analysis of the enzyme. European Journal of Biochemistry 184, 455–464.

Martin, D.P., Williamson, C., Posada, D., 2005. RDP2: recombination detection and analysis from sequence alignments. Bioinformatics 21, 260–262.

Miles, M.A., Souza, A., Povoa, M., Shaw, J.J., Lainson, R., Toye, P.J., 1978. Isozymic heterogeneity of *Trypanosoma cruzi* in the first autochthonous patients with Chagas' disease in Amazonian Brazil. Nature 272, 819–821.

Momen, H., 1999. Taxonomy of *Trypanosoma cruzi*: a commentary on characterization and nomenclature. Memorias Do Instituto Oswaldo Cruz 94, 181–184.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Souto, R.P., Fernandes, O., Macedo, A.M., Campbell, D.A., Zingales, B., 1996. DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. Molecular and Biochemical Parasitology 83, 141–152.

Sturm, N.R., Vargas, N.S., Westenberger, S.J., Zingales, B., Campbell, D.A., 2003. Evidence for multiple hybrid groups in *Trypanosoma cruzi*. International Journal of Parasitology 33, 269–279.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tibayrenc, M., 1995. Population genetics of parasitic protozoa and other microorganisms. Advances in Parasitology 36, 47–115.

Tibayrenc, M., 1998. Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. International Journal of Parasitology 28, 85–104.

Tibayrenc, M., Ayala, F.J., 1988. Isozyme variability in *Trypanosoma cruzi*, the agent of Chagas' disease: genetical, taxonomical, and epidemiological significance. Evolution 42, 277–292.

Tibayrenc, M., Cariou, M.L., Solignac, M., 1981. Genetic interpretation of flagellate zymograms of the genera *Trypanosoma* and *Leishmania*. Comptes rendus des seances de l'Academie des sciences III 292, 623–625.

Tibayrenc, M., Ward, P., Moya, A., Ayala, F.J., 1986. Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure. Proceedings of the National Academy of Sciences of the United States of America 83, 115–119.

Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends in Ecology and Evolution 11, 367–372.