Short communication

Differentiation of *Trypanosoma cruzi* I subgroups through characterization of cytochrome *b* gene sequences☆

Angel E. Spotorno O[a,*], Luis Córdova[a], Aldo Solari I[b]

[a] *Programas de Genética Humana, ICBM, Facultad de Medicina, Universidad de Chile, Independencia 1027, Clasificador No. 7, Casilla 70005, Santiago 7, Chile*
[b] *y Biología Celular y Molecular, ICBM, Facultad de Medicina, Universidad de Chile, Independencia 1027, Clasificador No. 7, Casilla 70005, Santiago 7, Chile*

A B S T R A C T

To identify and characterize chilean samples of *Trypanosoma cruzi* and their association with hosts, the first 516 bp of the mitochondrial cytochrome *b* gene were sequenced from eight biological samples, and phylogenetically compared with other known 20 American sequences. The molecular characterization of these 28 sequences in a maximum likelihood phylogram ($-\ln L = 1255.12$, tree length = 180, consistency index = 0.79) allowed the robust identification (bootstrap % >99) of three previously known discrete typing units (DTU): DTU IIb, IIa, and I. An apparently undescribed new sequence found in four new chilean samples was detected and designated as DTU Ib; they were separated by 24.7 differences, but robustly related (bootstrap % = 97 in 500 replicates) to those of DTU I by sharing 12 substitutions, among which four were nonsynonymous ones. Such new DTU Ib was also robust (bootstrap % = 100), and characterized by 10 unambiguous substitutions, with a single nonsynonymous G to T change at site 409. The fact that two of such new sequences were found in parasites from a chilean endemic caviomorph rodent, *Octodon degus*, and that they were closely related to the ancient DTU I suggested old origins and a long association to caviomorph hosts.

Keywords:
Cy*b*
Chile
Octodon degus
Caviomorpha
DTU I
DTU Ib

1. Introduction

Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease in humans, are found as parasites in different mammals and insects throughout the Americas. They exhibit considerable genetic heterogeneity (Miles et al., 1978; Tibayrenc et al., 1986), to the point that several subdivisions and names have been proposed according to the different methods applied (Momem, 2001). Nevertheless, two major groups have repeatedly emerged, being consensually referred as *T. cruzi* I and II (Anonymous, 1999); these can be further subdivided into six subgroups or discrete typing units (DTU) designated as DTU I, IIa, IIb, IIc, IId, and IIe (Brisse et al., 2001). Furthermore, the correlation between such subgroups and three maxicircle clades (Machado and Ayala, 2001) provided further distinctions: clade A corresponded to DTU I, clade C to DTU IIb, and clade B to the other strains (Brisse et al., 2003). Clade A is a single relatively homogeneous lineage, highly prevalent in ancient marsupial hosts in North America (Briones, 1999), South America (Yeo et al., 2005), and in the sylvatic triatomine *Mepraia spinolai* from Chile (Barnabe et al., 2001). By contrast, clade B has hybrid origins from representatives of DTU IIb and IIc (Machado and Ayala, 2001) (Westenberger et al., 2005). Some of these have distinct biological properties (Revollo et al., 1998), different infectivities and virulence in man (Breniere et al., 1998), and associations with different transmission cycles (Brisse et al., 2000).

We have recently collected 60 *T. cruzi* samples associated with different mammalian and insect hosts in North and Central Chile. In order to molecularly identify them and examine their attribution to previously detected subgroups (Brisse et al., 2003), we sequenced the 5′-half of the cytochrome *b* (Cy*b*) gene coding regions included within the maxicircle DNA of the mitochondrial genome in eight different samples. The molecular characterization of these sequences together with 20 available American sequences previously reported allowed the identification of all chilean samples and the detection of an undescribed new sequence found within an endemic chilean rodent. The relationships of this new molecular entity with one of the subgroups already described will be documented and discussed.

Sequences of the 573-nt fragments from the *T. cruzi* Tulahuen (Ochs et al., 1996) Cy*b* gene were obtained through the p18 (5′-GAC AGG ATT GAG AAG CGA GAG AG-3′) and p20 probes (5′-CAA ACC TAT CAC AAA AAG CAT CTG-3′). Amplifications were performed

☆ *Note*: Nucleotide sequence data reported in this paper are available in the GenBank database under the accesion numbers EU559323–EU559329.

* Corresponding author. Tel.: +56 2 978 6469; fax: +56 2 737 3158.
E-mail address: aspotorn@med.uchile.cl (A.E. Spotorno O).

during 35 cycles (94 °C, 1 min; 50 °C, 30 s; 72 °C, 90 s) followed by a final elongation step (7 min, 72 °C). Sequencing of PCR products was carried out using an automatic AbiPrism sequencer. Sequences were deposited in GenBank with accession numbers EU559323–EU559329.

Sequences were aligned using the programs Clustal X 1.8 (Thompson et al., 1997) included in MEGA 4 (Tamura et al., 2007), and by eye. The number of nucleotide substitutions and within and between groups diversity were obtained using MEGA4, and transition/transversion ratios estimated in PAUP* 4.0b8a (Swofford, 2002). Levels of genetic diversities were measured in terms of number of nucleotides (π per nucleotide site, i.e., the probability that two randomly chosen homologous nucleotides are different).

We searched first for the model of DNA substitution that best fitted the data using a hierarchical likelihood ratio test as implemented in the program Model Test 3.7 (Posada and Crandall, 1998). The model that best fitted these data was the HKY + G (a model of four free parameters with unequal base frequencies and inclusion of rate variation among sites, $-\ln L = 1569.96$, $G = 0.2255$). The individual and combined phylogenetic analyses were performed through PAUP 4.0b10 program using maximum-parsimony (MP) with the heuristic search option, neighbor-joining (NJ) (Saitou and Nei, 1987), and maximum likelihood (ML) methods. Character analysis and reconstruction of unambiguous character changes were performed through MacClade software (Maddison and Maddison, 1992).

The aligned matrix of Cyb gene sequences for all 28 samples showed 363 conserved characters, and 153 variable sites. Among these, 87 were phylogenetically informative. Mean base nucleotide compositions were 47.6% for T, 8.4% for C, 26.1% for A, and 17.9% for G, with those from individual sequences not substantially different from such averages. The number of base differences per sequence from mean diversity of the entire data set was 31.96 ± 3.21.

From the different analyses of the Cyb matrix of 516 bp, and particularly the maximum likelihood one, four robust terminal groups of *T. cruzi* sequences with bootstrap % >99 (Fig. 1) emerged, including a large heterogenous group (DTU IIc + d + e) weakly linked with bootstrap % = 69. The former were the same as three subgroups previously detected (Brisse et al., 2003): DTU IIb (clade C), DTU IIa, and DTU I (clade A), plus an undescribed sequence related to the latter. The number of base differences per sequence from mean intergroup diversity was 28.79 ± 3.03, but the number of differences within each subgroup ranged between 1 and 5.6 (Fig. 1).

DTU IIb (node 1 in Fig. 1) was the most divergent subgroup, being separated by 59.83 differences with DTU IIa (Fig. 1). It was characterized by 29 unambiguous nucleotide substitutions. Among these, four were nonsynonymous ones.

DTU IIa (node 2 in Fig. 1) was the least divergent lineage, being separated by 12.68 differences with the rest of clade B (node 3 in Fig. 1). It shared five synonymous substitutions, and was consistently linked within clade B. The latter also included one of the chilean samples (F06.V195.vinchuca) collected within peridomestic *Triatoma infestans*. All these formed the robust clade B (node 3 in Fig. 1) with a bootstrap % of 99, and was defined by seven unambiguous substitutions.

DTU I (node 4 in Fig. 1) was distinctive by its 38.78 differences with clade B, and by sharing five unique synonymous substitutions at sites 405, 420, 435, 483, and 489. This well robust subgroup (bootstrap % of 99) also now included the sequence from a sample found within the sylvatic chilean triatomine *M. spinolai* (sample A04.F. *M. spinolai*), in agreement with previous reports (Barnabe et al., 2001), as well as another found in the peridomestic mammal *Capra hircus* (sample 24t.p18.cabra).

An apparently new subgroup (node 5 in Fig. 1) linked four parasite sequences collected from particular chilean mammals. Two of them were found in wild *Octodon degus*, an endemic chilean caviomorph rodent (Woods, 2005), and the other two in peridomestic *C. hircus*. It was a robust (bootstrap 100%) and distinct subgroup, being characterized by 10 unique nucleotide

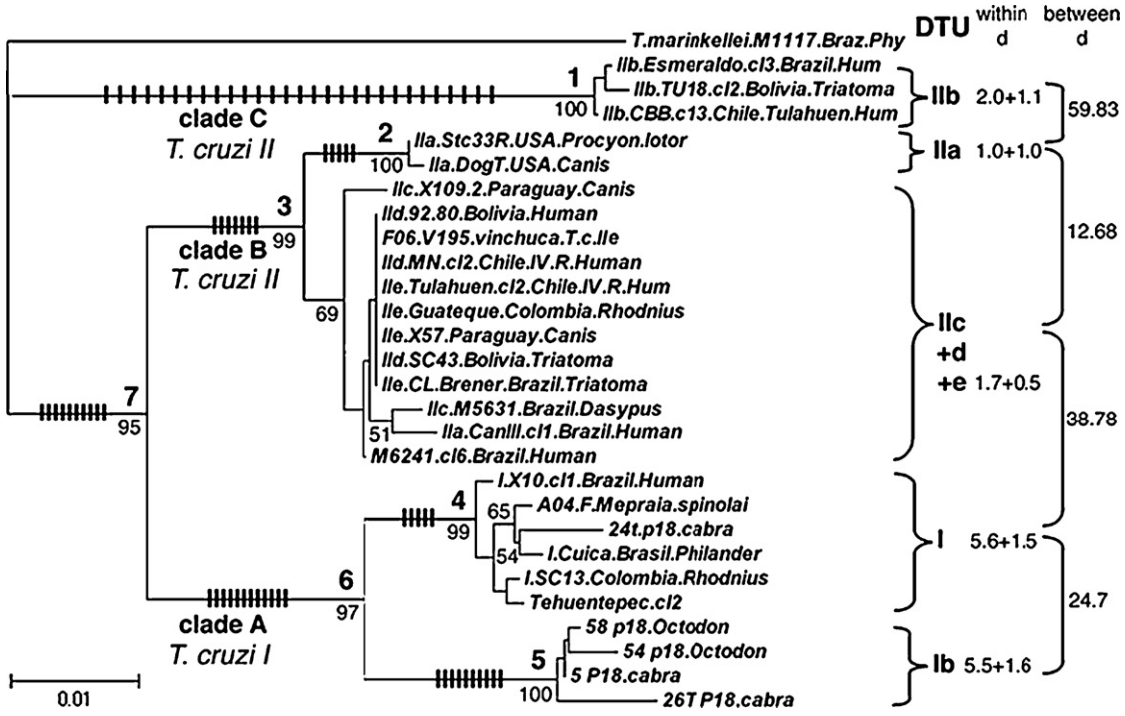

Fig. 1. Maximum likelihood phylogram of 28 *Trypanosoma cruzi* samples reconstructed from mitochondrial Cyb sequences. Tree of $-\ln L = 1255.12$, tree length = 180, consistency index = 0.79, recalculated index = 0.93. Nodes marked with numbers; bootstrap percentages obtained after 500 replicates included under each node. Each bar in branches marks one unambiguous nucleotide substitution. Number of differences (d) in nucleotides within and between subgroups added to the right.

substitutions, among which a single one was a nonsynonymous G to T change at site 409.

Such undescribed subgroup (node 5) and the well-known DTU I sequences (node 4) were consistently linked by means of node 6 (Fig. 1), with a bootstrap % of 97. Both were separated by 24.7 nucleotide differences. This extended clade A shared 12 unique substitutions, among which four were nonsynonymous ones. Given the close phylogenetic relationships, the new subgroup might be designated as DTU Ib, if further investigations with other molecular markers and widespread samples support this subdivision.

The phylogenetic analysis of all available Cy*b* sequences thus allows a robust definition and unambiguous characterization of most *T. cruzi* clades and subgroups previously described, and the identification and distinction of all chilean samples. These include clades C and A, which also were the most divergent in all trees based on COII, Cy*b* and ND1 sequences, as well as with microsatellite data (de Freitas et al., 2006), and finally the apparently new DTU Ib. This finding, together with the comparison of two complete maxicircle sequences (Westenberger et al., 2006), and the very recent identification of four new *T. cruzi* I haplotypes from Colombian samples (Herrera et al., 2007) indicate that there is still much genetic variation to uncover in the complex sylvatic cycles of this parasite. Moreover, additional molecular markers should be used to distinguish some subgroups unresolved through Cyb variation, like DTU IIc + d + e (Fig. 1).

The clade A subgroups seem to be among the oldest *Trypanosoma* lineages. The split between DTU I and clade B (node 7) was estimated with Cy*b* data as 4.4 ± 1 million years (Myr), based on maximum likelihood lengths, no rejection of a clock-like behavior, and assuming a 1% substitution rate per Myr (Brisse et al., 2003). If we assume that such estimate is also true for our topology (Fig. 1), the split between DTU I and the new DTU Ib (node 6 in Fig. 1) should be around 2.7 Myr.

In addition to such old origins, both clade A subgroups are preferential parasites of the oldest South American mammals, marsupials and caviomorphs (Briones, 1999; Spotorno and Walker, 2000; Yeo et al., 2005). For instance, DTU I was the preferential lineage found in 8 out of 10 marsupial species (Yeo et al., 2005), but it was also the most common parasite found in six out of nine caviomorph species included in their Table 3. Similarly, the newly detected DTU Ib was also found within *O. degus*, an endemic caviomorph rodent from Central Chile, whose specific separation from other species of the genus have been dated at 2.57 + 0.76 Myr, based on independent molecular estimates (Opazo, 2005), and calibrated with the appearance of the first caviomorph fossil in America (Wyss et al., 1993).

Such preferential parasite–host associations appear to be the product of significant biological facts, as shown in experimental infections, where *T. cruzi* I promoted permanent and very mild infections in marsupials, in contrast to those by *T. cruzi* II, which were eliminated in similar marsupial infections (Briones, 1999). The finding that both *T. cruzi* I subgroups have been found in association with caviomorph rodents corroborates their close phylogenetic relationships. This seems to be another case of the "ecological host fitting" hypothesis, postulated as the principal mechanism in the evolution of trypanosomes (Hamilton et al., 2007). Nevertheless, the fact that both *T. cruzi* I subgroups were also able to infect modern domestic mammals like goats, indicates the high adaptability of this widespread multihost parasite.

Acknowledgement

This work was supported by a Fondo Nacional de Ciencia y Tecnología grant (FONDECYT 1040762, Chile, to A. Solari).

References

Anonymous, 1999. Recommendations from a satellite meeting. Memorias Do Instituto Oswaldo Cruz 94, 429–432.
Barnabe, C., Neubauer, K., Solari, A., Tibayrenc, M., 2001. *Trypanosoma cruzi*: presence of the two major phylogenetic lineages and of several lesser discrete typing units (DTUs) in Chile and Paraguay. Acta Trop. 78, 127–137.
Breniere, S.F., Morochi, W., Bosseno, M.F., Ordonez, J., Gutierrez, T., Vargas, F., Yaksic, N., Noireau, F., 1998. *Trypanosoma cruzi* genotypes associated with domestic *Triatoma sordida* in Bolivia. Acta Trop. 71, 269–283.
Briones, M.R.S., 1999. The evolution of two *Trypanosoma cruzi* subgroups inferred from rRNA genes can be correlated with the interchange of American mammalian faunas in the Cenozoic and has implications to pathogenicity and host specificity. Mol. Biochem. Parasitol. 104, 219–232.
Brisse, S., Barnabe, C., Tibayrenc, M., 2000. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. Int. J. Parasitol. 30, 35–44.
Brisse, S., Verhoef, J., Tibayrenc, M., 2001. Characterisation of large and small subunit rRNA and mini-exon genes further supports the distinction of six *Trypanosoma cruzi* lineages. Int. J. Parasitol. 31, 1218–1226.
Brisse, S., Henriksson, J., Barnabe, C., Douzery, E.J., Berkvens, D., Serrano, M.D, Carvalho, M.R., Buck, G.A., Dujardin, J.C., Tibayrenc, M., February 2, 2003. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. Infect. Genet. Evol. 2, 173–183.
de Freitas, J., Augusto-Pinto, L., Pimenta, J.R., Bastos-Rodrigues, L., Gonzalves, V.F., Teixeira, S.M.R., Chiari, E., Junqueira, A.C.V., Fernandes, O., Macedo, A.M., Machado, C.R., Pena, S.D.J., 2006. Ancestral genomes, sex, and the population structure of *Trypanosoma cruzi*. PLoS Pathog 2, 0226–0235.
Hamilton, P.B., Gibson, W., Stevens, J.R., 2007. Patterns of co-evolution between trypanosomes and their hosts deduced from ribosomal RNA and protein-coding gene phylogenies. Mol. Phylogenet. Evol. 44, 15–25.
Herrera, C., Bargues, M.D., Fajardo, A., Montilla, M., Triana, O., Vallejo, G.A., Guhl, F., 2007. Identifying four *Trypanosoma cruzi* I isolate haplotypes from different geographic regions in Colombia. Infect. Genet. Evol. 7, 535–539.
Machado, C.A., Ayala, F.J., 2001. Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. Proc. Natl. Acad. Sci. U.S.A. 98, 7396–7401.
Maddison, W.P., Maddison, D.R., 1992. MacClade v.3: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA.
Miles, M.A., Souza, A., Povoa, M., Shaw, J.J., Lainson, R., Tpye, J.J., 1978. Isozymic heterogeneity of *Trypanosoma cruzi* in the first autochthonous patients with Chagas disease in Amazonian Brazil. Nature 272, 819–821.
Momem, H., 2001. Some current problems in the systematics of Trypanosomatids. Int. J. Parasitol. 31, 640–642.
Ochs, D.E., Otsu, K., Teixeira, S.M.R., Moser, D.R., Kirchhoff, L.V., 1996. Maxicircle genomic organization and editing of an ATPase subunit 6 RNA in *Trypanosoma cruzi*. Mol. Biochem. Parasitol. 76, 267–278.
Opazo, J., 2005. A molecular timesacale for caviomorph rodents (*Mammalia hystricognathi*). Mol. Phylogenet. Evol. 37, 932–937.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Revollo, S., Oury, B., Laurent, J.P., Barnabe, C., Quesney, V., Carriere, V., Noel, S., Tibayrenc, M., 1998. *Trypanosoma cruzi*: impact of clonal evolution of the parasite on its biological and medical properties. Exp. Parasitol. 89, 30–39.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Spotorno, A.E., Walker, L.I., 2000. Origen y evolución de los mamíferos chilenos. In: Muñoz, A., Yañez, J. (Eds.), Mamíferos de Chile. Ediciones CEA, Valdivia, Chile, pp. 217–227.
Swofford, D., 2002. PAUP* (Phylogenetic Analysis Using Parsimony). Sinauer Associates, Sunderland, MA.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4 0. Mol. Biol. Evol. 24, 1596–1599.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1997. The ClustalX-Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
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. Proc. Natl. Acad. Sci. U.S.A. 83, 115–119.
Westenberger, S.J., Barnabe, C., Campbell, D.A., Sturm, N.R., 2005. Two hybridization events define the population structure of *Trypanosoma cruzi*. Genetics 171, 527–543.
Westenberger, S.J., Cerqueira, G.C., El-Sayed, N.M., Zingales, B., Campbell, D.A., Sturm, N.R., 2006. *Trypanosoma cruzi* mitochondrial maxicircles display species- and strain-specific variation and a conserved element in the non-coding region. BMC Genomics 7, Article number 60.
Woods, C., 2005. Infraorder hystricognathi. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal Species of the World: A Taxonomic and Geographic Reference, vol. 2. The John Hopkins University Press, Baltimore, MD, pp. 1538–1600.
Wyss, A.R., Flynn, J.J., Norell, M.A., Swisher, C.C., Charrier, R., Novacek, M.J., McKenna, M.C., 1993. South-America earliest rodent and recognition of a new interval of mammalian evolution. Nature 365, 434–437.
Yeo, M., Acosta, N., et al., 2005. Origins of Chagas disease: *Didelphis* species are natural hosts of *Trypanosoma cruzi* I and armadillos hosts of *Trypanosoma cruzi* II, including hybrids. Int. J. for Parasitol. 35, 225–233.