Cluster Analysis of Non-conserved Proteins of *Trypanosoma cruzi* Reference Strains Displays Parity between these Groupings (Peptidemes) and the Consensually Accepted Parasite Lineages

Felipe S. COELHO¹, Danielle P. VIEIRA², Angela H. LOPES¹, Maria A. SOUSA³

¹Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, RJ, Brasil
²Polo Avançado da Universidade Federal do Rio de Janeiro, Macaé, RJ, Brasil
³Laboratório de Toxoplasmose e Outras Protozooses, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, RJ, Brasil

Abstract. The protein profiles of the epimastigote stages from eight reference strains of *Trypanosoma cruzi* belonging to three different lineages (TcI, TcII and TcVI) were analyzed by SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis), under standardized conditions. More than 40 protein bands were observed in each strain. Around 55% of them were not shared by all stocks (non-conserved proteins), representing their intra-specific variability. Then, they were coded for processing by numerical taxonomy, using three association coefficients and the UPGMA clustering algorithm. With all coefficients assayed, two major groups were clearly seen, confirming the dichotomy within *T. cruzi* taxon, as demonstrated by other molecular and biochemical approaches. In the present study, the term peptideme was used to name the groups of strains based on their polypeptide profiles, following the above-cited methodology. Then, two major peptidemes were identified, each one presenting subdivisions. The isolates identified as TcI clustered in the same major peptideme, displaying a subgroup with the opossum isolates (G, SC28, Dm28c) apart from the stock of human origin (Colombian strain). The other major peptideme also showed two subgroups, regardless the coefficient used. One of them included the TcII strains (Y, SF21), both from Brazilian patients, and the other the TcVI stocks, both originally from triatomines from Southern Brazil (CL Brener, FL). As far we know, this is the first report on the parity between the *T. cruzi* lineages consensually accepted and their grouping into peptidemes based on SDS-PAGE and the numerical analysis of non-conserved proteins.

Key words: *Trypanosoma cruzi* lineages, reference strains, SDS-PAGE analysis, numerical taxonomy, peptidemes

INTRODUCTION

*Trypanosoma cruzi* Chagas, 1909 (Protozoa: Kinetoplastea, Trypanosomatidae) is the etiological agent of Chagas disease (American trypanosomiasis), mainly in Latin American countries, where it infects 6–7 millions of people (WHO 2019). In the last decades, *T. cruzi* infections have spread around the world due to the mobility of asymptomatic patients from endemic areas. Nowadays, Chagas disease has been considered an emergent global health problem (Tanowitz *et al.* 2011, WHO 2019). This parasite presents high intra-specific diversity and may cause distinct clinical forms. Accordingly, the typing of *T. cruzi* samples is of utmost importance (Zingales *et al.* 2009, 2012; Consentino and Agüero 2012).
The identification of *T. cruzi* can be very easy, since its typical forms could be seen in Giemsa-stained preparations (×1,000) (Sousa 1999, Oliveira et al. 2017). Otherwise, the intra-specific variability of *T. cruzi* can also be evaluated using classical morphological and biological approaches (e.g. Brener 1965, Andrade et al. 1974, Sousa 1999, Kikuchi et al. 2010). However, the increasing development of biochemical and molecular techniques in the study of *T. cruzi* has substantially improved our knowledge on its identification, besides genomic and proteomic diversity. At present, at least six lineages (TcI–TcVI), previously named DTUs (discrete typing units) (Brisse et al. 2009, 2012). Another *T. cruzi* genotype (Tcbat) was described by Marcili et al. (2009), and has been confirmed by other authors (e.g. Consentino and Agüero 2012, Lima et al. 2015). These genetic types can be identified using several techniques, among them, the classical isoenzyme analyses (Miles et al. 1978), which gave us the first irrefutable evidences on the six DTUs in *T. cruzi* (Tabayrenc and Ayala 1988). Several genomic techniques have been employed to identify and to type *T. cruzi* stocks, among them the analyses of the restriction fragment length polymorphism (RFLP), followed by molecular hybridizations (Morel et al. 1980, Macedo et al. 1993). It was of utmost importance the implantation of the polymerase chain reaction (PCR) (Sturm et al. 1989, Fernandes et al. 2001) for identification and characterization of *T. cruzi* strains, as well as for diagnostic use. For improving the PCR results, the products can be subsequently sequenced, treated with restriction enzymes (PCR-RFLP), besides submitted to other amplifications, sometimes with several primers (e.g. Brito et al. 2008, Luna-Marin et al. 2009, Consentino and Agüero 2012, Zingales et al. 2012). Modified versions of the PCR, using combinations with other technologies, have also allowed valuable contributions to the clinical and epidemiological researches, as well as to basic studies (Cura et al. 2015). However, the majority of the above-cited techniques can be very expensive and usually restricted to specialized laboratories. There are also several biochemical techniques for identification of proteins expressed by distinct living beings. Some of them are very complex and applied in proteomic analyses, usually addressed to questions of clinical interest. However, in the last decades, proteomic analyses have also been used in more advanced studies for characterization and typing of several trypanosomatid species (e.g. El-Sayed et al. 2005, Kikuchi et al. 2010, Avila et al. 2016, Oliveira et al. 2018).

The analysis of the whole-proteins using one dimensional (1 D) sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli 1970) is a very simple and low expensive analytical technique with a good resolving power (Gibson et al. 1978, Janssen and Van Bijsterfeld 1981, Rosa et al. 2000). SDS-PAGE with adaptations, or associated to Western blotting, has been frequently used in study of trypanosomatids to identify specific proteins or antigens, searching for answers to biological and taxonomic questions, as well as to improve the diagnosis of organisms of medical interest (e.g. Contreras et al. 1985; Pinho and Giovanni-de-Simone 1989; Saldaña et al. 1993, 1998; Branquinha et al. 1995; Mejía et al. 2004; Ziccardi et al. 2005; Añez-Rojas et al. 2006; Gomes et al. 2008; Moraes et al. 2008; Cervantes-Landin et al. 2014).

Taylor and Williams (1977) and Taylor et al. (1982, 1983) used SDS-PAGE for distinguishing some *Trypanosome* species and their variants, including *T. cruzi*. Taylor et al. (1983) presented an excellent visualization of the SDS-PAGE profiles of three *T. cruzi* stocks belonging to different zymodemes (nowadays identified as TcI, II and IV), thus displaying protein bands with possible diagnostic value. However, in the same study, they analyzed several other *T. cruzi* isolates, and using subjective criteria proposed their classification into five peptidemes (subpopulations having similar polypeptide profiles), which displayed an incomprehensible correlation with their zymodemes.

In the present study, the whole-protein extracts of epimastigotes (in exponential growing phase) from eight reference strains of *T. cruzi* belonging to TcI, TcII and TcVI lineages, were analyzed by 1 D SDS-PAGE. As this technique generates complex banding patterns, we selected those protein bands not shared by all stocks (not-conserved proteins) to be coded and analyzed by numerical taxonomy procedures (Sneath and Sokal 1962). Interestingly, we could evidence strong correlations between the strain groups based on protein profiles (peptidemes) and their known lineages [DTUs, zymodemes (Z)], according to the consensually accepted nomenclature (Zingales et al. 2009).

**MATERIALS AND METHODS**

Eight reference stocks of *T. cruzi* were used in the present work. These samples were from the trypanosomatid collection maintained by one of the authors (M. A. Sousa) at the Oswaldo Cruz Institute. They were identified by their names, code-numbers, original hosts and geographical origins, besides their phenotypic and genotypic classifications (see Table 1).
The analysis of T. cruzi strains were performed using (1-D) SDS-PAGE based on Laemmli (1970). Before analysis, all stocks had been maintained under standardized culture conditions, as follows. They were passed three times, at six-day intervals, in liver infusion tryptose (LIT) medium supplemented with brain heart infusion (BHI) and 10% fetal calf serum. Thereafter, exponential growing cells from 6 day-old cultures (epimastigote stages) were harvested and washed twice by centrifugation (1,000 × g, 10 min) in a phosphate-buffered saline (PBS), pH 7.2. The pellets were frozen in liquid nitrogen in a buffer containing 20 mM Tris-HCl at pH 7.2. After twice freezing and defrosting, the cell pellets were resuspended in a lysis buffer (20 mM HCl, 15 mM NaCl, 1 mM Na3VO4, 1 mM EDTA, 1 mM EGTA, 10 mM NaF, 0.02% sodium azide; pH 8.0) with protease inhibitors and 1% SDS (pH 7.2). Subsequently, the suspensions were homogenized and centrifuged at 5,000 × g for 10 min (4°C). Whole-cell protein extracts from the supernatant of each strain were dosed according to the Lowry method (Lowry et al. 1951). Aliquots containing 60 μg of the proteins from each stock were separated in 10% SDS-PAGE at 250 V and 25 mA, for 90 min, using a 16 cm high system (GE Healthcare Life Sciences). Molecular mass markers were also included in the gel. Coomassie brilliant blue was used for staining and visualization of the protein patterns. Those bands not shared by all stocks (twenty-two non-conserved proteins) were regarded the most representative of the variability in T. cruzi, then being selected for analysis according to numerical taxonomy principles (Sneth and Sokal 1962).

A data matrix was constructed with the eight strains of T. cruzi, these being “operational taxonomic units” (OTUs), and their selected bands, as “character units”, with the records of presence (1) or absence (0) of each one per strain. The NTSYS software (version 1.7) was employed for data processing, using three association coefficients [Jaccard, Dice and Simple Matching (SM)] and the UPGMA (unweighted pair-group method using arithmetic averages) clustering algorithm. This was followed by the construction of phenograms, which represented the global similarity among the stocks analyzed.

RESULTS AND DISCUSSION

In our experimental conditions, complex protein patterns were found in all T. cruzi stocks analyzed by SDS-PAGE, each one displaying more than 40 protein bands. All isolates shared ∼25% of these bands (conserved proteins), including those with molecular mass compatible to heat shock proteins (HSPs) and the glycoprotein gp63 (Fig. 1). All stocks displayed bands varying between >170–17 kDa, the more prominent ones ranging from ∼95–72 kDa, >55–34 kDa, and 17 kDa. They appear typical of T. cruzi, since they were also found in other studies (Taylor et al. 1983, Cervantes-Landin et al. 2014). Otherwise, nearly 55% of the protein bands were not shared by all stocks (non-conserved proteins), thus being the most useful data to analyze the intra-specific variability in T. cruzi. These proteins were lightly stained, but could be visualized and were coded for numerical analysis. They are diagrammatically represented in Fig. 2, where those bands exclusive of some strains were also highlighted with rectangles.

Trypanosoma cruzi SDS-PAGE banding patterns of the present study could not be easily correlated with those from other studies, where whole-cell proteins of epimastigotes from axenic cultures of this species were analyzed (Taylor et al. 1983, Barr et al. 1990, Mejia et al. 2004). This probably was due to operational differences, such as the polyacrylamide concentrations, electrophoretic running conditions and gel sizes, since they could affect the resolution of the protein bands, as previously observed (Taylor et al. 1982). However, we
could identify several conserved proteins of *T. cruzi*, as seen in our SDS-PAGE analysis, in the gels from the study of Cervantes-Landín et al. (2014).

The “deme” terminology was proposed by Hoare (1967) to be used as name suffix of subgroups within a species or subspecies, according to any chosen criterion (*e.g.* zymodeme, schizodeme). Taylor et al. (1982) proposed the name “peptideme” for grouping species of *Trypanosoma (Schizotrypanum)* from bats based on the visualization of some distinctive proteins in SDS-PAGE, then describing eight peptidemes. However, in that study, distinct species were placed into a same peptideme, whereas strains of a same species were classified into distinct peptidemes. Subsequently, Taylor et al. (1983) also analyzed three *T. cruzi* reference stocks from distinct zymodemes (Z1, Z2 and Z3, currently corresponding to TcI, TcII and TcIV) by SDS-PAGE. They obtained an excellent gel (see p. 357) in which some protein bands perfectly distinguish those stocks, probably being of diagnostic value. In the same study, the authors analyzed thirty-six *T. cruzi* isolates whose zymodeme had been previously determined, only one being reference stock (TcI), and proposed their classification into five peptidemes. However, the dispersion of the zymodemes within each peptideme remained barely understandable. This was probably due to the visual choice of the set of protein bands used to delimit each peptideme, besides the use of a single reference strain.

Aiming to a better understanding of the complex banding patterns generated by SDS-PAGE in *T. cruzi* (Fig. 1), we used a more objective procedure for grouping the stocks analyzed in the present study, as that provided by numerical taxonomy (Sneath and Sokal 1962). It is worthy mentioning that the numerical taxonomy has been successfully used following SDS-PAGE, mainly with bacteria and fungi, disclosing reliable groupings (Costas 1990, Rosa et al. 2000, Rodrigues et al. 2004).

In the present work, using the three association coefficients, the UPGMA phenograms displayed two major groups of strains in *T. cruzi*, which were identified herein as major peptidemes, and referred as **mP 1** and **mP 2** (Fig. 3). This confirmed the dichotomy within the *T. cruzi* taxon, as previously reported by several authors (Tibayrenc et al. 1993, Souto et al. 1996, Brisse et al. 2000). The numbering of these major peptidemes followed that used by Brisse et al. (2000) for the two major *T. cruzi* lineages (1 and 2) (Figs 2, 3). Both **mP 1** and **mP 2** also presented subgroups, these herein referred simply as peptidemes (P), also confirming the diversity within each major *T. cruzi* lineage (Tibayrenc et al. 1993, Brisse et al. 2000).

In our study, the **mP 1** only included strains identified as TcI (formerly DTU 1, Z1). However, using the SM coefficient, which considers positive and negative matches, a clear subdivision was observed (Fig. 3). One of them only included isolates from opossums (G, SC28, Dm28c), remaining apart a strain of human origin from Colombian. In the last decade, studies with different molecular markers showed a remarkable genetic diversity within TcI. Several authors proposed its subdivision into five subtypes, identified as TcFa–TcFe, whereas others recommended to group them only in two genotypes, one being domestic (identified as Tc_Dom) and the other sylvatic (León et al. 2015). However, at present, no consensual nomenclature for the TcI subtypes was defined (Ramírez and Hernández 2018). Therefore, in the present study, all strains from the **mP 1** remained classified into a single peptideme, herein identified as **P I**.

As illustrated in Fig. 3, **mP 2** displayed two well distinct subgroups/peptidemes (P), regardless the association coefficients used. One of them, herein identified as **P II**, included two strains of human origin from Brazil (Y and SF21), both previously classified as TcII (formerly DTU 2b, Z2). The other, identified as **P VI**, presented two stocks from triatomine bugs from Southern Brazil (CL Brener and FL), both belonging to TcVI (formerly DTU 2e, ZB). It is worthy mentioning that to avoid confused interpretations, the numbering of the peptidemes described in the present work (P I, P II, P VI) is the same of their genetic types (TcI, TcII and TcVI), as consensually defined (Zingales et al. 2009).

CONCLUSIONS

Our study indicates that the analysis of non-conserved proteins of *T. cruzi* stocks, as seen in SDS-PAGE and treated by numerical taxonomy procedures, can arise consistent intra-specific classifications, as reported for other microorganisms (*e.g.* Costas 1990, Rosa et al. 2000, Rodrigues et al. 2004). It was evidenced that *T. cruzi* strains belonging to the same genetic type (formerly DTU, zymodeme) were clustered, in a parallel manner, into a same peptideme. This demonstrates that the protein profiles from SDS-PAGE are also bona fide markers, as the isoenzymes, which provided the pioneering identification of the six *T. cruzi* lineages (Tibayrenc and Ayala 1988, Brisse et al. 2000), then...
Fig. 1. Total protein profiles of eight *Trypanosoma cruzi* reference strains separated in 10% SDS-PAGE at 250 V, 25 mA, 90 min, and stained by Coomassie brilliant blue. The position of some conserved proteins is indicated on the right. M: molecular mass markers. (kDa) are indicated on the left.
evidencing correlations between phenotypes and genotypes. Accordingly, we suggest that the present study be expanded to include analyses of a larger number of *Trypanosoma* reference strains with distinct genotypes, as well as other *Trypanosoma* species, mainly to identify protein bands with diagnostic value. Finally, considering that SDS-PAGE analysis has low cost, fast and easy execution, using a single gel, it can be also used as a simple alternative technique for routine controlling of the authenticity of trypanosomatids maintained in cultures, since the procedure can be performed under rigorously standardized conditions.

**REFERENCES**


Cluster Analysis of Proteins in T. cruzi


Received on 1st June, 2019; revised on 25th September, 2019; accepted on 4th November, 2019