Hitchhiking *Trypanosoma cruzi* minicircle DNA affects gene expression in human host cells via LINE-1 retrotransposon

Augusto Simões-Barbosa, Enrique R Argañaraz, Ana Maria Barros, Ana de Cássia Rosa, Nivaldo P Alves*, Patrícia Louvandini, Marian R D'Souza-Ault, Nadjar Nitz, Nancy R Sturm**, Rubens J Nascimento, Antonio RL Teixeira/+

Laboratório Multidisciplinar de Pesquisa em Doença de Chagas, Faculdade de Medicina, Universidade de Brasília, Caixa Postal 04536, 70919-970 Brasília, DF, Brasil* Tecnogene Diagnóstico, Brasília, DF, Brasil **Department of Immunology, Microbiology and Molecular Biology, David Geffen School of Medicine, University of California at Los Angeles, US

The horizontal transfer of Trypanosoma cruzi *mitochondrial minicircle DNA to the genomes of naturally infected humans may play an important role in the pathogenesis of Chagas disease. Minicircle integrations within LINE-1 elements create the potential for foreign DNA mobility within the host genome via the machinery associated with this retrotransposon. Here we document integration of minicircle DNA fragments in clonal human macrophage cell lines and their mobilization over time. The movement of an integration event in a clonal transfected cell line was tracked at three months and three years post-infection. The minicircle sequence integrated into a LINE-1 retrotransposon; one such foreign fragment subsequently relocated to another genomic location in association with associated LINE-1 elements. The p15 locus was altered at three years as a direct effect of minicircle/LINE-1 acquisition, resulting in elimination of p15 mRNA. Here we show for the first time a molecular pathology stemming from mobilization of a kDNA/LINE-1 mutation. These genomic changes and detected transcript variations are consistent with our hypothesis that minicircle integration is a causal component of parasite-independent, autoimmune-driven lesions seen in the heart and other target tissues associated with Chagas disease.*

Key words: kinetoplast - horizontal transfer - p15 - mutagenesis - pathogenesis - retrotransposition

Trypanosomes are members of the Order Kinetoplastida and represent a highly diverse and specialized group of eukaryotic organisms. *Trypanosoma cruzi* is transmitted by triatomines (Hemiptera: Reduviidae) to mammals; the infection occurs by invasion of phagocyte cells at the entry site (Teixeira et al. 2006). The infection can be transmitted by blood transfusion or transplacenta from mother to offspring. The migration of *T. cruzi*-infected individuals to North America, Europe, and Asia poses a threat to countries outside the normal range of the parasite, as these individuals are potential blood donors. Chagas disease affects 18 million people in Latin America and is characterized clinically by a spectrum of manifestations often seen several decades after establishment of the initial infection (WHO 2002). Chagas disease is the most lethal endemic human parasitic disease in the Western Hemisphere. It has been estimated that Chagas heart disease kills 100,000 people and accounts for 6 billion dollars worth of disability-adjusted life per year (WHO 1996). There is no vaccine against Chagas disease and its treatment is presently unsatisfactory

(Lauria-Pires et al. 2000). The pathologic lesions found in the heart and digestive system of chronic Chagas patients may not appear for decades after the initial infection, and may not be explained exclusively by active parasite tissue destruction. The minimal rejection unit, consisting of lyses of parasite-free target host cells by the immune system mononuclear cells, is a common denominator of pathology in Chagas disease (Teixeira et al. 2006).

T. cruzi is characterized by the presence of a single mitochondrion with topologically interlocked double-stranded maxicircles and minicircles that constitute almost 15% of the cellular DNA, representing the largest amount of extra-nuclear genetic material in any cell (Lukes et al. 2002, Liu et al. 2005, Junqueira et al. 2005). This kinetoplast DNA network is composed of several dozen maxicircles that code for two large ribosomal RNAs and a hydrophobic subset of mitochondrial proteins (Westenberger et al. 2006), along with several thousand minicircles that provide the information for the post-transcriptional process of RNA editing in the form of guide RNAs (Avila & Simpson 1995, Simpson et al. 2004). Each minicircle possesses four 122-bp conserved regions interspersed by four variable regions that contain guide RNA genes, totaling approximately 1400 bps (Sturm et al. 1989).

Our recent studies focused on the integration of *T. cruzi* minicircle fragments into host LINE-1 retrotransposons associated with chagasic-like pathology in the chicken and rabbit models, and in human patients showing heart arrhythmias characteristic of Chagas disease (Nitz et al. 2004). This work extended previous observations suggestive of minicircle integration into the

host genome (Teixeira et al. 1994) with important implications for presentation of the autoimmune aspect of Chagas pathology in humans, resulting in publication in the prestigious journal *Cell*. Questions arose regarding authenticity of the integration site data, leading to a unilateral retraction by the *Cell* Editor (Marcus 2005). No experimental data contradicting the Nitz et al. (2004) paper was presented, nor was the implied contamination of the Chagas patients' DNA samples articulated. The validity of the reported rabbit and chicken model systems was not in question. It is our hope that this ominous Retraction will be rescinded by the *Cell* Editorial Board. We stand by our data, and continue to pursue the logical consequences of the phenomenon described.

The estimated 1,000,000 copies of mobile LINE-1 elements are considered a primary force in vertebrate genome evolution over the last 150 million years (Smit et al. 1995, Ostertag & Kazazian 2001). The LINE-1 transposable elements of the family Ta is approximately 6 kb. Full length, active LINEs are capable of retro-transposing to different sites in the genome. Most LINEs are truncated subsets, and a vast majority of copies shorter than 0.5 kb (Pavlícek et al. 2002a) are present in the human genome in proximity to A+T repeat-rich (G-banding in the chromosomes) regions, constituting preferential exon shuffling sites (Ostertag & Kazazian 2001). Furthermore, LINEs are free-loaders of repeat-rich SINEs (IAPs, HERVs, MIRs, and Alu-like microsatellites) generating local polymorphisms (Pavlícek et al. 2002b). Full-length LINEs have a poorly-characterized internal promoter at their 5'-end open reading frame; ORF-1 encodes a nucleic acid binding p40 protein; ORF-2 encodes a protein with endonuclease, reverse transcriptase, and zinc finger-like motif (Feng et al. 1996). The LINE-1 short 3'-untranslated region ends in a polyadenylated (poly A) tail that is important for reverse transcription; this region carries typical structural characteristics of genomic LINE-1s, such as inversions, deletions, and 5'-end microhomologies to the target DNA sequence (Babushok et al. 2006). Active LINE-1 retro-transposition involves functional elements employing a 'copy and paste' mechanism by which the retransposon proteins assemble with the RNA from which they were translated (Kazazian 2000). Genome analysis has identified an estimated 61 active LINE-1 elements in human cells (International Human Genome Sequencing Consortium 2001).

In view of the demonstrable preference for integration of minicircles within LINE-1 elements (Nitz et al. 2004), we postulated that the foreign insertions could be mobilized within the host genome. The functional impact of host genome alterations could include changes in endogenous gene expression as well as the generation of chimeric host-minicircle gene products, either of which may contribute to chronic Chagas disease. Due to the frequency and complexity of the phenomenon, examination of whole organisms is impractical; thus, we have adapted macrophage cell culture for the infection of *T. cruzi* and subsequent characterization of parasite DNA integration, mobilization and phenotypic effect. Here we describe the acquisition and movement of integrated parasite DNA within the human macrophage genome in culture and document a spe-

cific alteration in gene expression over time. The variable factor of time in this process may explain the decades of delay experienced between initial infection and displays of pathogenesis.

MATERIALS AND METHODS

Macrophage and parasite growth and infection - The human macrophage (ATCC number: CRL-1593.2) line U937 and the *T. cruzi* were grown respectively as described (Nitz et al. 2004). Trypomastigotes were used to infect macrophages in culture at a ratio of 5:1. At this ratio the phagocytes eradicated the infection within two weeks, as demonstrated by co-culture in liver infusion tryptose medium and by infection of weanling mice.

Authentication of the macrophage U937 line - Genotypic characterization of the U937 cell was achieved by analyses of autosomal short tandem repeats (STRs) THO.1 and TPOX, and of AMELY (Amelogenin) *loci*. Genetic profiles were determined by simultaneous polymerase chain reaction (PCR) amplification of STRs and sex linked amelogenin, as described in the ATCC Cell Biology Catalog. The amplification products were separated in a 4% polyacrylamide/8 M urea and were analyzed by silver staining according with standard procedures (Murkerjee & Biswas 2005).

Primers and probes used - The primer pair sequences used for authentication of the U937 *Homo sapiens* macrophage DNA samples were as follows:

1) TPOX forward primer 5'-ACTGGCACAGAACAGGCACTTAGG-3', and reverse primer 5'-GGAGGAACTGGGAACCACACAGGT-3' (Beckmann & Weber 1992).

2) THO1 forward primer 5'-ATTCAAAGGGTATCTGGGCTCTGG-3', and reverse primer 5'-GTGGGCTGAAAAGCTCCCGATTAT -3' (Bell et al. 2000).

3) AMELY forward primer 5'-GACCAGAATATGAGACAGGAACTG-3', and reverse primer 5'- TTGCTAAGTTAAGTGATTGTATAAGCA-3' (Lukka et al. 2005).

The *T. cruzi* kDNA specific primer sets Sk35/36 and Sk34/67 (Sturm et al. 1989), *T. cruzi* nuclear DNA specific Tcz1/2 (Moser et al. 1992) and the kCR probe are cited in previous papers (Nitz et al. 2004). The clones representing PCR products were cloned from template DNA amplifications from post-infection parasite-free macrophages.

1) S34 (5'-ACACCAACCCCAATCGAACC)/ S67 (-GGTTTTGGGAGGGGCGTTCA).

2) S35 (5'-ATAATGTACGGG(T/G)GAGATGC)/ S36 (-GGTTCGATTGGGGTTGGTG).

3) Tcz/1 (5'-CGAGTCCTTGCCCACAACGGG)/ Tcz2 (-CCTCCAAGCAGCGGATAGTTC-3').

4) L1 represents a PCR amplification of clone A (GenBank: AF002199) sequence obtained with antisense primer L1 (51-GCCCATTCAGTATGATATTGG) and the primer SP6 included in the TA vector (Invitrogen). This product was cleaved with *Hinf*I yielding a 280 bp DNA stretch that was used as probe.

5) kCR (5'-TTTTGGTTTTGGGAGGGGCGTTC AAATTTTGGCCCGAAAATTCATGCATCTCCCCCGTACATT ATTTGGCCGAAAATGGGGGGTTGTTCGATGGA GGTGAGGTTCGATTGGGGTTGGTGTAAG) represents a constant region of *T. cruzi* Berenice kDNA minicircle showing primers S34, and S67 in both extremities, and the nested S35 antisense primer from bp 65 to 46 (underlines). This kCR sequence was a PCR amplification product from *T. cruzi*-infected macrophage DNA template with primers S34/S67, which was cloned in the TA vector. The probes were ^{32}P-dATP-labeled using the Random Primer DNA Labeling Kit (Invitrogen) according to the manufacturer's protocol.

Cloning and sequencing PCR amplifications of the p15 cDNA and gene was achieved with kCR probe specific for minicircles and Va 470 bp probe (GenBank accession number AY589192) specific for p15 gene, and primers: *VSP6/5* 5'-AGC GAG GCA GAG GTT TTT GG 3, *VSP4* 5'-GTT GGC CAC AGC AGG CAA G 3', *VSP3r* 5'-TCA CCT TCT GAA TCC ATC CA-3', *VSP2r* 5'-GCC AGC TGA AGG AAC AGA T 3', 5'-AGG TTT GAT ATG GCT CAG 3'. Combinations of primer pairs *VSP6/5* with *VSP1r* and *VSP4* with *VSP3r* and subsequently with *VSP2r* were used for each amplification reaction aiming at reproducible demonstration of kDNA insertion. The relative minicircle and p15 gene primer locations are shown in Fig. 6. Amplification products were cloned into the PGEM-T (Promega) and commercially sequenced.

Cloning post-infection macrophages - This investigation was conducted by inoculation of *T. cruzi* into a monolayer of U937 cells growing in 15 ml tissue culture flask. Clones were obtained from a single post-infection macrophage by end point dilution. Adherent human peritoneal macrophages (10^4) were used as a feeder-layer disseminated into 96-well plates. The plates were kept in a 5% CO_2 atmosphere and 95% humidity. Each clone derived from *T. cruzi*-infected cell culture was subsequently expanded in 24 wells plate for DNA extraction. The selection of each cell clone presenting the kDNA integration was done by PCR with specific primers set. The screening showed $52 \pm 8\%$ kDNA-positive clones in multiple experiments.

In situ hybridization - The spreads of U937 macrophages that had undergone *T. cruzi* infections were obtained after 3 h co-culture in the presence of 20 μl of a 1 M colchicine. The interphase and metaphase chromosomes were collected on glass slides and processed for in situ hybridization (Teixeira et al. 1994). After hybridization with L1 probe the metaphase chromosome spreads were soaked in 0.1 M phosphate buffered solution (PBS, pH 7), containing Triton X-100 to 1% (v/v) final concentration; dehybridization was secured by incubation of the metaphase plate glass slide in the PBS-Triton solution for 1 h at 37ºC, after which it was washed trice in PBS and dipped in a mix of 50% formamide with 2X SSC, at 92ºC, for 5 min. Probe removal was confirmed by microscopic exam of the glass slide showing absence of previous fluorescent spots using a Zeiss Axiophot microscope with an HBO-50 ultraviolet light source. Then the glass slide was rehybridized with the biotin-labeled kCR probe (10 ng

μl) overnight in a moist chamber at 42ºC. After three washes in 50% formamide in 2X SSC, the spreads were flooded with a 1:250 dilution of rabbit anti-biotin antibody, and incubated for 45 min at 37ºC. Again, the slides were washed three times and incubated with a 1:40 dilution of fluorescein-conjugated goat anti-rabbit IgG. The meta-phase nuclear elements were counterstained with 1% propidium iodide solution and the slides were analyzed under the microscope with ultraviolet light (Teixeira et al. 1994).

Nucleic acid analyses - DNA was extracted from the following sources: 1) uninfected macrophage U937 line; 2) archetype Berenice strain *T. cruzi*; 3) *Homo sapiens* blood mononuclear cells; 4) BALB/c mouse spleen cells. The rational for choice of the restriction enzyme was the pattern of cuts in the archetype (Thertulien et al. 1994) minicircle sequence: *Eco*R1, single cut; *Nsi*I single cut in each out of four conserved regions of a minicircle; *Bam*HI and *Xho*I producing digestion at all. PCR was conducted with specific primers for parasite kDNA and nuclear DNA, and for host DNA sequences defined in the text. The protocols for RT-PCR used in the differential display of mRNA used random primers with 1 μg of template RNA (Liang & Pardee 1992).

RNA was purified using TriZOL reagent (Invitrogen), fractionated in agarose gels, and blotted onto charged nylon membranes (ClonTech). Hybridizations were performed as recommended by the manufacturer. The membrane was washed once for 5 min with 1X SSC and 0.1% SDS, twice for 30 min each with 0.2X SSC and 0.1% SDS, and autoradiographed for variable periods of time. cDNA fragments from recombinant plasmid clones were used as probes on RNA blots.

Detection and cloning of differentially expressed transcripts - Differential display of mRNA from kDNA-positive and from kDNA-free macrophages was used (Liang & Pardee 1992). DNA-free total RNA was subjected to RT-PCR by combining one- and two-bases anchored primers with random primer sets HT11C/HAP57 and HT11C/HAP60 (GenHunter Corporation). The candidate cDNA bands, identified in triplicate reactions, were eluted from the gels, and their 3' fragments were reamplified and cloned. The differential expression of each transcript was analyzed by hybridization to total cell RNA/or poly A selected RNA rapid amplification of 5' cDNA (5' RACE, Invitrogen) was used to obtain the 5' end of the cDNA.

RESULTS

Horizontal transfer of minicircles via living infection in culture - In order to better follow the integration of minicircles in the human genome and to track subsequent changes in location over time as postulated due to their association with mobile genetic elements, a protocol using cultured macrophages was implemented. Live *T. cruzi* cells were used to infect an immortalized macrophage population, followed by recloning and extended maintenance of the parasite-free lines in cell culture.

The DNA from U937 macrophage cultures infected with *T. cruzi* was subjected to Southern hybridizations with a homologous minicircle conserved region (kCR) probe.

The DNA from the *T. cruzi*-infected cells showed a distinct band pattern that was not present in the *T. cruzi* DNA tested with the kCR probe, thus indicating that the minicircle DNA is in a different configuration in the infected macrophage DNA sample compared to DNA from the kinetoplast. The *T. cruzi*-infected macrophage samples taken on the 7th post-infection day showed different band patterns; in addition to a 360-bp band representing the kinetoplast, larger size bands of 1.2, 1.8, and 2.2 kb were formed with the macrophages DNA, which were not present in the parasite DNA. At the 30th day post-infection the lower band representing the *T. cruzi* kDNA was absent in the host macrophages, demonstrating that living parasites had been eradicated (Fig. 1A). Control experiments showed no hybridization with the protozoan nuclear DNA probe or with a cloned maxicircle gene DNA probe (a gift from David Campbell, UCLA). These results were confirmed by PCR amplification of minicircle sequences present in DNA from *T. cruzi* alone and from macrophages infected with *T. cruzi* using kDNA primer set S34/67. Furthermore, absence of amplification of the 0.2-kb band from the *T. cruzi* nuclear DNA primer set Tcz1/2 (Moser et al. 1992), which is a highly repetitive sequence comprising 9% of total nuclear DNA, and, also, absence of a 0.36-kb band in the Southern blot, both with DNA

collected on the 30th day post-infection macrophages appeared to correlate with persisting kDNA integration in host cells that eradicated the living infection (Fig. 1B, C).

The absence of the 0.2-kb minicircle-derived band in the Southern blot and the absence of the 0.2-kb parasite nuclear DNA PCR products in the 30th day post-infection macrophage DNA indicate that the *T. cruzi* infection has been eradicated, leaving minicircle sequences with altered mobilities integrated in the host cell genome. Transfer of minicircle sequences from *T. cruzi* into the macrophage genome occurred as a natural consequence of infection.

Co-localizing minicircle integration with LINE-1 in whole chromosomes - In order to confirm integration of the kDNA insertion in LINE-1 elements, in situ hybridization was carried out utilizing L1 and kCR probes and metaphase plates of macrophages that were infected with *T. cruzi*.

kDNA co-localized in LINE elements of two separate chromosomes, as both probes hybridize to the same region (Fig. 2A, B). Control experiments showed that neither the kCR nor the L1 probes hybridized, respectively, with human nuclear DNA or with *T. cruzi* kDNA in metaphase plates of non-infected macrophage chromosomes (data not shown). Thus, the integration event was maintained within this cell line. It must be noted that the LINE-1 probe would be predicted by interact with thousands of sites within the genome, but that in control cell lines no signal is seen (data not shown). This is a notable feature of LINEs, with interspersed direct and inverse SINE repeats generating DNA condensation normally making them inaccessible to probes (Cassavant et al. 2000). In the case of the G10 macrophage cell line, the presence of the kDNA integration prevented this condensation, allowing the detection of a single LINE among the multitude.

The co-localization of minicircle and LINE-1 element sequences provides further evidence for the reported association (Nitz et al. 2004), and portends the coming observations of minicircle mobility as a consequence of that association.

Fig. 1: integration of *Trypanosoma cruzi* minicircles into the genome of macrophages as a consequence of infection. A: Southern hybridization of *Nsi*I digests of 7-day infection macrophages and of 30-day (post-infection) macrophages with kCR probe on blots of 0.8% agarose gels; B: ethidium bromide staining of minicircle PCR amplification products using primers S34/67. Profiles formed with *T. cruzi* and with 7-day infection macrophage are altered in post-infection macrophages by the absence of 0.12 kb band and its catamers; C: ethidium bromide staining of nDNA PCR amplification products using Tcz1/2. The absence of the 0.36 kb band in the Southern blot and of the 0.2 kb PCR products in the 30th day post-infection macrophage DNA indicates that the *T. cruzi* infection has been eradicated, leaving the integrated minicircles in the host cell genome.

Fig. 2: co-localizing *Trypanosoma cruzi* kDNA minicircle sequences into chromosomes of metaphase plate post-infection macrophage. A: LINE-1 showing fluorescent lights up in two metaphase chromosomes probed with a specific L1 probe (see methods); B: co-localizing kDNA minicircle sequences within LINE-1 with kDNA specific kCR probe.

Amplification of integrated sequences captures associated host LINE-1 elements - Macrophage cultures undergoing continuous replication for 3 months after the living *T. cruzi* infection was eradicated were used to obtain DNA at the integration junctions. These samples represented our starting point for the characterization of minicircle mobilization within the host macrophage genome. We sought to characterize the integrations using direct blotting techniques in combination with PCR to define the sites of integration.

PCR amplification products were generated using this post-infection macrophage DNA as template with the kDNA primer set S35/36 (Sturm et al. 1989). No amplification was obtained from template DNA employing parasite-specific nuclear DNA primer pairs, corresponding to ribosomal RNA and a *T. cruzi* spliced leader RNA genes (data not shown). Preliminary attempts to clone a restriction fragment representing 2.2-kb kDNA integration into post-infection macrophages were unsuccessful; cloning in competent cells yielded truncated sequences, usually smaller than 0.5 kb, representing rearranged fragments of kDNA and no flank host DNA. The difficulties observed with TA-vector transformed SURE cells could not be resolved in any other recombinase-free *E. coli* strains tested. The sequencing results from some DH5α rearranged clones frequently showed the S36 primer in direct and inverse repeats, suggesting that linearized minicircle insertions in the macrophage genome occurred in a direct repeat or inverse manner, and identifying associated SINEs with frequently rearranged highly-interspersed repeat sequences, thus forming palindromes that creates the substrate for deletion and recombination by the host cell (data not shown). These results and the experience of others (Sligtom et al. 1992, Gilbert et al. 2002) were used to design a PCR approach to determine the region(s) of minicircle integration.

We succeeded in capturing the initial integration events using PCR amplification by employing only one specific primer, S36 against a motif in the minicircle conserved region, for amplification. Amplification products were cloned into the PCRII vector and screened with *T. cruzi* kDNA probe, resulting in the sequence of five representative clones (A to E), ranging from 527 to 700 bp. These samples presented the LINE-1 ORF2 segment between outer kDNA regions not showing kink repeat motifs that prevent cloning and sequencing. Each clone contained minicircle and LINE-1 DNA sequences (GenBank numbers AF002199 to AF002203) loaded with SINEs (HERV, MIRs, and Alu-like), suggesting that the minicircle insertions had occurred within these highly repetitive elements. Fig. 3 illustrates the integration of kDNA minicircle sequences in LINE-1 AC134888.2. In four clones (A, C, D, and E) minicircle constant and variable regions were flanked by mammalian Alu sequences followed by LINE-1 DNA. These clones showed a homologous segment to human chromosomes Y, 4, and 13; clones A, C, and E

Clone C

```
  1   GGTTCGATTG GGGTTGGTGT AATTAGGAGG AGTTGAGTGG AAGGGAAATG GAGTTGAGTG
 61   AAGGATGGAA TGGAATGAGA TGGAATGGAA TGGAATGGAA TGGAATGGAA TGGAATGGAA
121   TGGAATGGAA TGGAATGGAA AGGGAAGGTA AGGGTGGTG G GAAGTAAGGT GGTGCGCTAA
181   TCCTCAATAA AATACTACCA AACCGATCCA GCAGCAAATC AAAAAGCTTA TCCACCATGA
241   TCCAAACTGG GCTGCATCCC TGGGGTCGAA GGCTGGTTCA CATCACAATC AATCAATCAC
301   ATATCCCAAA CCCACGTCCA AAACTCAATG ATATGTCAAT GATGCAGAAA GCCTTTTGAC
361   AAAATTCAAC AACCTTCATT GCCTAAAACT CTCAATAAAT TAGATATTGA TGGGATTGGT
421   ATCTCAAGGG ATGTATCTCT AATATTAAGA GTCATTTAGT CAAACCCACA GCCAATATCA
481   TACTGAATGG GTNAAAAACC TGGAAGGCAT TCCCTTTGAA AATCTTCAGT TACCTTTCTT
541   TTGGATATGT GTGTGAGGGA AAGAGTACCT ACAAACTACC CTTTTAGTTA AATGCCATAT
601   ACAATAAATA TTAATACCTT ACATGTCCTC ATGTTGTATA TTACACCAAC CCCAATCGAA
661   CC
```

Fig. 3: schematic representation of *Trypanosoma cruzi* kDNA minicircle sequences integration in human LINE-1 retrotransposon. Truncated fragments of kDNA minicircles found in the LINE-1 copy at chromosome Y; clone C showing conserved region primer S36 (dark blue) at both ends followed by minicircle variable regions (light blue) and the LINE-1 sequence (green).

showed the S36 primer followed by the sequence AATTAG at the end of the constant region. The B clone contained minicircle sequence at only one end showing the S36 primer followed by the sequence AATATATG. In clones A, C, D, and E the minicircle conserved regions were in tandem inverse orientation consistent with two insertions in opposite orientations. Clone B appears to represent priming of S36 from *T. cruzi*-infected macrophage DNA and may represent outer junction sequences. BLASTn analyses showed LINE-1 in each case, showing average identity of 93% (clones A, D, and E, respectively GBAC147649.2, AC069302.4 and AC096740.3), and of 92% (clones B and C, respectively AL160255.14 and AC134882.2); clones A and C targeted chromosome Y; whereas clones D and E derived from chromosome 4 and clone B inserted in a chromosome 13 LINE-1 copy. BLASTn extended results also revealed that clones A and E include regions with 88% identity with satellite MIR II, and CpG putative island in clone A with 95% identity. On 3 occasions (clones A, C, and D) the juxtaposition of kDNA with LINE-1 sequences were truncated by microsatellite DNA with evident microhomology, suggesting homologous recombination mediated minicircle insertions in LINE-1 within the human host cells.

A natural extension of this analysis is the search for potential chimeric proteins or sites of recombination associated with the integration events. BLASTx searches detected ORF-2 segments in clones A, B, and E, showing similarity with viral endonuclease: human reverse transcriptase (hRT). BLASTx also revealed homologies with ubiquitin like protein (clone B), and with hypothetical chimera human and *T. cruzi* proteins (clones A, C, D, and E). Thus, kDNA insertions were found in five out of five instances within a transposon of the LINE-1 Ta family, revealing a hot spot for minicircle integration.

Mobilization of the minicircle integration - To document any movement of the integrated minicircle we cloned post-infection macrophages and maintained them in active culture for an extended period of time. Genomic DNA from these lines was analyzed for shifts in minicircle distribution by Southern hybridization. Changes in minicircle distribution can then be localized for additional effects.

Macrophage clonal line G10 showed positive hybridization with the kCR probe and was kept in continuous culture for three years with an average doubling time of 16 ± 2 h. G10 genomic DNA digested with *Eco*RI or with *Xho*I displayed distinct band patterns at three years post-cloning versus three months (Fig. 4), with pattern alterations consistent with minicircle movement. The 6.4-kb band indicated that the integrated kDNA was mobilized from one site to another with a loss of repetitive elements rather than being dispersed throughout the G10 genome.

This result indicates that our hypothesis was correct, in that integrated foreign sequences can be mobilized within the host genome. The foundation of this prediction is that the minicircle sequence was relocated by virtue of association with active retrotransposons, however direct recombination within the genome could also explain the observation.
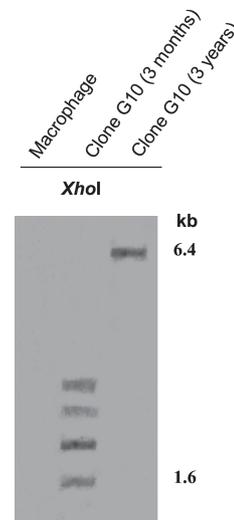
Fig. 4: alterations in localization associated with mobilization of a chimeric minicircle-LINE-1 sequence. *Xho*I digestions of genomic DNA from macrophages were separated in 0.8% agarose gels blotted and kCR probed. Note marked variations in profiles of kDNA positive bands from G10 DNA at three months and three years post-infection.

Alteration detected in three-year-old culture - Minicircle transposition within the macrophage genome prompted us to investigate the possibility of genetic changes in the macrophage clone G10. Functional consequences of minicircle integration will be manifest in detectable phenotypes.

A profile of mRNAs from clonal lines maintained for 3 years was compared with that of the kDNA-free control line using differential display (Liang & Pardee 1992). An endogenous mRNA was under-expressed in 3-year-old G10 (Fig. 5A), and identified subsequently from the uninfected sample as a conserved transcript showing identity with *Mus musculus* Sub1 mRNA (GenBank: NM011294) and with *H. sapiens* p15 cDNA encoding an activated RNA polymerase II transcription factor at chromosome locus 5p13.3 (GenBank: BC022339.1). Note that the stretch of LINE-1 mobilizing kDNA inserted in the p15 gene showed identity with *H. sapiens*, and that samples of U-937 macrophage line used in this study were re-authenticated by molecular tests (Lukka et al. 2005, Mukherjee & Biswas 2005); mouse DNA template was used as a negative control in authentication (data not shown). Hybridization of RNA from three-year-old G10 culture with a p15-specific probe confirmed its absence as a stable mRNA species (Fig. 5B), a transcriptional alteration exclusive to the three-year-old G10 cell line. Control hybridization with actin probe confirmed equivalent loading of RNA (Fig. 5C). Southern hybridization of G10 DNA with the p15 probe confirmed that the p15 gene was modified at three years post-infection versus three months post-infection and control DNA (Fig. 5D).

The minicircle movement correlated with a reduction in expression of an endogenous host gene in the clonal cell line. The identification of the gene allowed demon-
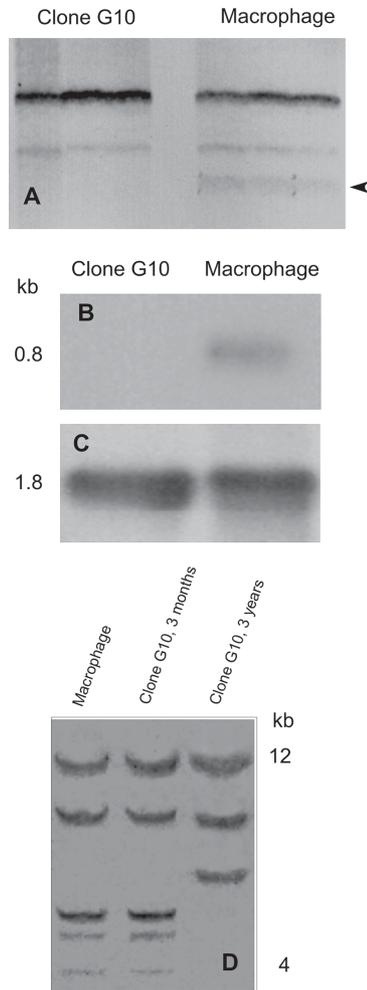
Fig. 5: transcript modification based on minicircle mobilization. A: differential display and mRNA analysis in G10, three years post-infection. Lack of a band is indicated by arrow; B: RNA hybridization of a 0.8-kb mRNA species with p15 cDNA probe; C: the 1.8-kb band represents the β-actin positive control; D: Southern hybridization of *Bam*HI digestion of DNA from control macrophages and the G10 with the p15 cDNA probe in the blot of a 0.8% agarose gel. Notice alteration of band profiles in the 3 years post-infection macrophage sample.

stration that the change can be seen in the DNA profile by Southern analysis; the three-month-sample of the G10 clonal line is identical to the uninfected macrophage profile, implying that the starting sample was unaffected for the expression of this gene. Our next goal was to look for an association of the affected product's gene with minicircle-derived sequence.

Mutagenesis of a host gene by minicircle transposition - The straightforward explanation for the p15 effect is a direct interaction with the mobilized minicircle fragment, whereby the foreign sequence has acted as a mutagen or has disrupted the cognate p15 promoter. The rearrangement of the p15 locus and loss of p15 transcript could be unrelated to the movement of minicircle sequence, caused by independent recombination events, secondary effects of the minicircle relocation, or senescence of the cell line.

The genetic link between p15 expression and minicircle integration was determined using PCR with various combinations of minicircle- and p15-specific primers (Fig. 6A), exposing an insertion of LINE-1 and minicircle sequence in chromosome locus 5p13.3 (GenBank accession numbers AY584192 and AY584193), 262-bp downstream from the p15 initiation codon. The chimeric DNA insertion in the G10 macrophage line (Fig. 6A, B) was cloned and found to contain two truncated minicircle conserved and variable regions linked to an active host LINE-1. Databank analysis revealed the truncated LINE-1 to show 99% identity (expect = 0.0) to human chromosome 4 *locus* 4g21 (GenBank: AC004061.1). However, a limited search showed that the 301–855-bp stretch has 92% identity (expect = 0.0) to the human β-globin frame locus at chromosome 11 (Gi: 28380636; NG: 000007.3), indicative of other sequences acquired by the element while in transit. The presence of a premature translation termination codon in the chimeric ORF at aa 132 could be the cause of transcript destabilization (Fukuda et al. 2003).

The footprints of the minicircle's travels through the host genome were captured by the associated LINE-1 characteristics. The recovered sequences indicate that the original minicircle integration may have relocated to the p15 location on chromosome 5 by hitchhiking along with a LINE-1 element originating on chromosome 4. The documentation of this event serves as an indicator of what is happening on the broader scales of decades of incubation and the array of cell types present in the host organism.

DISCUSSION

This work documents a unique but representative genome modification resulting from *T. cruzi* minicircle integration and mobilization of the integrated DNA by the action of an active LINE-1 retrotransposon in a cultured line of human macrophage. In order to confirm the kDNA integration into LINE-1 elements probes for minicircle and LINE-1 sequences in host cell genomes were used. The kDNA insertion in LINE-1 elements was co-localized in the metaphase plate of a post-infection macrophage by in situ hybridization. A minicircle fragment linked to the truncated LINE-1 element's most 3'-end poly-A tail inserted in the open reading frame of the p15 gene on chromosome 5, locus 5p13.3; a stretch of sequence homology with a retrotransposon on chromosome 4 indicated the path of this minicircle-loaded element. The presence of a specific β-globin cluster signature within the truncated LINE-1 suggests that the kDNA could have mobilized via chromosome 11 to chromosome 5. Genomic DNA from several macrophage lines was assayed by Southern hybridization, revealing multiple instances of kDNA integration. Minicircle sequence elements were found in five out of five occasions within Alu-like SINE-associated LINE-1 elements.

The LINE-1 transposable elements with high-interspersed repeat sequences were rearranged frequently in the macrophage chromosome, showing local polymorphisms. Such kinked DNA regions have prevented the extension of sequencing towards both arms in order to study nearby genome organization, a common difficulty

(Babushok et al. 2006). Similar unique features have been associated with genome instability and non-clonability (Razin et al. 2001, Venter et al. 2001, Symer et al. 2002, Pavlicek et al. 2002b). These issues were compounded by the integrated minicircles, whose redundant conserved regions are interspersed by variable regions showing Alu-like and microsatellite repeats.

The p15 gene derived from the human macrophage cell line yielded 99% identity (expect = 2e-143) with *M. musculus* and 85% identity (expect = 5e-46) with *H. sapiens* at the 5' end, and, respectively 98% and 94% identities at the 3' end (GenBank: BC010967.1 and U12979.1). The LINE-1 mobilizing the minicircle sequence was 99% homologous (expect = 0.0) with *H.*
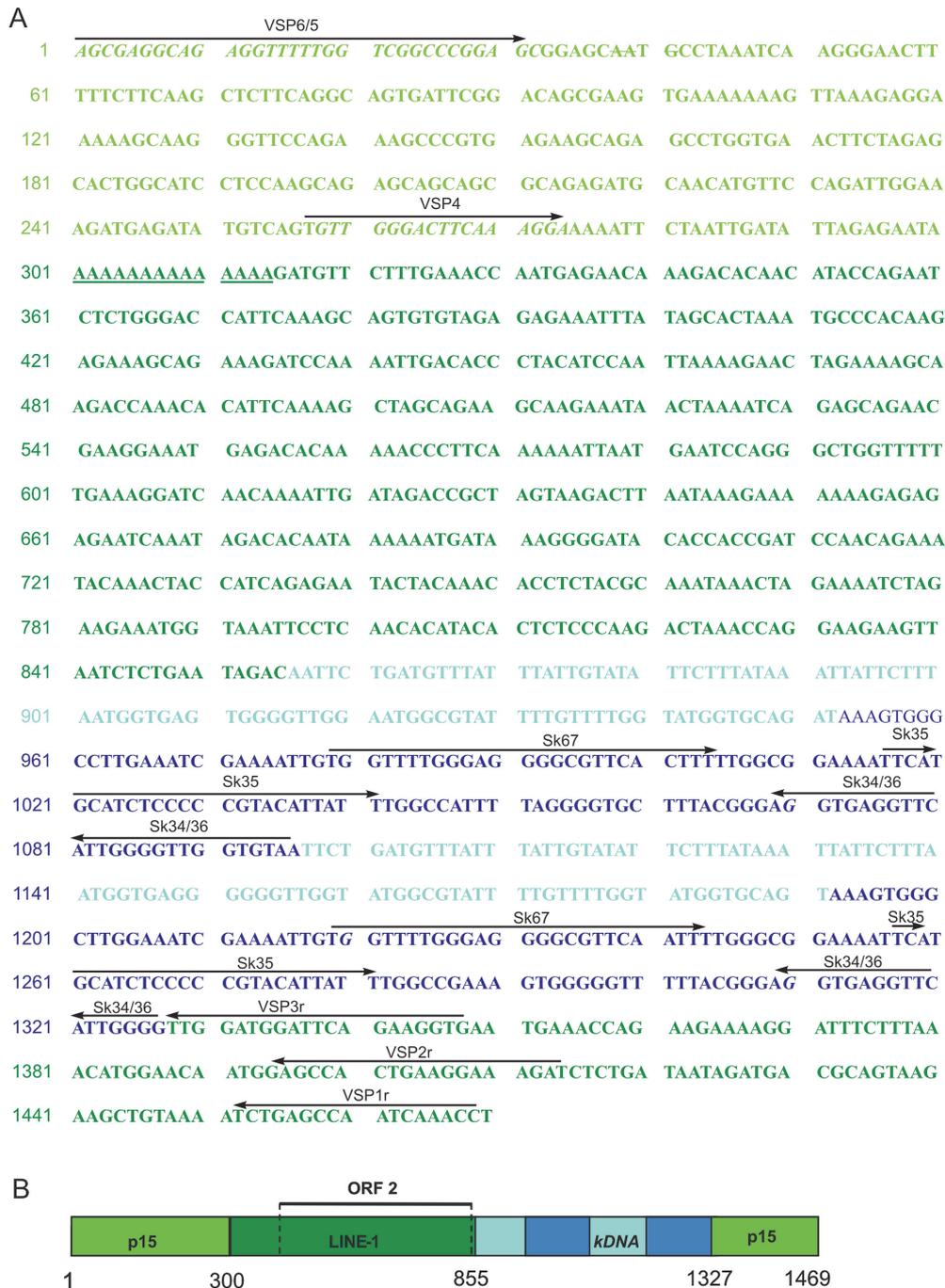
Fig. 6: modification of the p15 gene by the insertion of minicircle kDNA linked to an active LINE-1 at chromosome 5p13.3. A: rupture of the p15 ORF by the insertion of truncated LINE-1 mobilizing a minicircle fragment. Sequence in light green is from the p15 gene, dark green from LINE-1; light blue shows the minicircle variable regions, and dark blue the minicircle conserved region. The p15 ORF (initiation codon bp 39-41) is interrupted at nt 300 by the LINE poly-A tail to bp 855, and extends to the minicircle bp 1327. The p15 gene continues from 1328 bp towards its 3' end. The primers used are indicated by arrows; B: schematic representation of the chimeric minicircle-LINE-1 showing insertion in the ORF of the p15 gene, and ORF2 similar to hRT from nt 410 to 851.

sapiens. SNPs and gaps accounting for polymorphisms present in the human macrophage p15 sequences could explain the differences in identity levels, hence the variance of the cloned U937 DNA. Furthermore, the p15 gene copy ruptured by truncated minicircle and LINE-1 could be amplified from G10 DNA template with other gene-specific primers designed from different regions at either end.

It is reasonable to expect that retrotransposons have been subjected randomly to DNA replication associated errors over time, thus accumulating SNPs and other mutations resulting from deletions and rearrangements, particularly in direct and inverted repeat regions at the 3' end of LINEs. In addition, the phenomenon of horizontal gene transfer (HGT) resulting from integration of minicircle sequences into the host cell genome only adds to the DNA instability at their insertion site. The finding of several SNPs and a three base pair deletion in the host cell sequence flanking the under-expressed p15 gene may explain why it shows a higher similarity to *M. musculus* than to *H. sapiens*. This type of sequence variation occurs in instances of HGT and usually requires the recalibration of the molecular clock (Simonson et al. 2005).

Polymorphism at sites of kDNA integration in the host cell has been documented with the identification of several dozen sequences originating from our studies and accessible through GenBank. The human genome is filled (Human Genome Sequencing Consortium 2001) with repeat sequences, including interspersed repeats derived from transposable elements (LINEs carrying SINEs on their backs), and long genomic regions that duplicate in tandem, palindromic or dispersed fashion. These duplicated segments create deletions during recombination (International Human Genome Sequence Consortium 2001). New approaches such as primer walking tail-PCR (Babushok et al. 2006) along with classic techniques such as genomic libraries in cosmid and lambda vectors or even BAC libraries will be pursued to circumvent the recombination problems inherent in characterization of the minicircle integration events; screening host DNA with minicircle probes will select the fragments carrying the chimeric hallmarks. These strategies should circumvent the current problems with PCR amplification and cloning.

Understanding of Chagas disease has been approached through correlations of phylogenetics and pathology (Campbell et al. 2004). Nuclear markers define 6 interrelated groups or discrete typing units (Westenberger et al. 2005), subdividing the 2 major groups of *T. cruzi* (Anonymous 1999); maxicircle markers identify 3 clades (Machado & Ayala 2001). Minicircles, a fluid population due to the functional flexibility of guide RNAs, have yet to reveal useful markers for epidemiology (Junqueira et al. 2005). The interplay between pathogen genotype and the minicircle integration that is a component of *T. cruzi* infection may reveal itself in the frequency of the integration event for particular subgroups or parasite populations. The variables of invasion, establishment within the host cell, and stability will all factor into the success of an individual parasite genotype. The relative composition of the heterogeneous minicircle population will also play a critical role in determining the mutagenic capacity of a given parasite.

The acquisition of foreign DNA and its persistence in progeny is a major finding that may result in gene loss or differential regulation. This phenomenon offers a novel strategy for mutagenesis: Ubiquitous host LINE-1 elements provide the basic retrotransposon machinery, while kDNA serves as both a tag and a secondary mutagen, thus defining the birth of another gene transfer tool that is both interspecific (parasite to host) and intraspecific (within the genome). In the case described by macrophage line G10, the p15 transcript appeared to be destabilized; the inhibition of transcriptional elongation by LINE-1 might profoundly affect expression of endogenous human genes (Kadlec et al. 2004, Han et al. 2005). However, an alternative outcome of integration and mobilization is the generation of a novel chimeric protein (Nitz et al. 2004, Inácio et al. 2004).

Similar phenotypic changes resulting from kDNA-driven mutagenesis may trigger the autoimmune rejection of parasite-free target host cells that are the pathologic substrate of the lesions present in the heart and the digestive system of patients that die of Chagas disease (Teixeira et al. 2006). These symptoms present 20 to 30 years after the initial infection; this delay may reflect the time required to accumulation of mutagenic transposition and following autoimmune response. Considering the complexity of the entire host environment, and the approximately 61 active LINE-1 elements in the average human host (International Human Genome Sequencing Consortium 2001), multiple LINE-1/kDNA insertions could accumulate to culminate in the autoimmunity characteristic of Chagas patients in the absence of persistent parasitemia. A study has been initiated for the detection of genome modifications resulting from kDNA integration into LINE-1 that correlate with various manifestations of Chagas disease.

REFERENCES

Anonymous 1999. Recommendations from a satellite meeting. *Mem Inst Oswaldo Cruz 94*: 429-432.

Avila HA, Simpson L 1995. Organization and complexity of minicircle-encoded guide RNAs in *Trypanosoma cruzi*. *RNA 1*: 939-947.

Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian Jr HH 2006. L1 integration in a transgenic mouse model. *Genome Res 16*: 240-250.

Beckmann JS, Webber JL 1992. Survey of human and rat micorssatellites. *Genomics 12*: 627-631.

Bell B, Nievas P, Abecia E, Martinez-Jarreta B, Hinojal R, Martinez-Cordero A 2000. Population genetics of the STR loci HUMCSFIPO. HUMF13A01, HUMFES/FPS and D12391 in Asturias (Northern Spain). *Forensic Sci Internl 113*: 21-23.

Brisse S, Henriksson J, Barnabé C, Douzery EJ, Berkvens D, Serrano M, De Carvalho MR, Buck GA, Dujardin JC, Tibayrenc M 2003. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. *Infect Genet Evol 2*: 173-183.

Campbell DA, Westenberger SJ, Sturm NR 2004. The determinants of Chagas disease: connecting parasite and host genetics. *Curr Mol Med 4*: 549-562.

Cassavant NC, Scott L, Cantrell MA, Wiggins LE, Baker RJ, Wichman HA 2000. The end of the LINE?: Lack of recent L1 activity in a group of South American rodents. *Genetics 154*: 1809-1817.

Feng Q, Moran J, Kazazian H, Boeke JD 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell 87*: 905-916.

Fukuda A, Tokonabe S, Hamada M, Matsumoto M, Tsukui T, Nogi Y, Hisatake K 2003. Alleviation of PC4-mediated transcriptional repression by the ERCC3 helicase activity of general transcription factor TFHH. *J Biol Chem 278*: 14827-14831.

Gilbert N, Lutz-Prigge S, Moran JV 2002. Genomic deletions created upon LINE-1 retrotransposition. *Cell 110*: 315-325.

Han JS, Szak ST, Boeke JD 2005. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes *Nature 429*: 270-274.

Inácio A, Silva AL, Pinto J, Ji Xx, Morgado A, Almeida F, Faustino P, Lavinha J, Liebhaber SA, Romao L 2004. Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay. *J Biol Chem 279*: 32170-32180.

International Human Genome Sequencing Consortium 2001. Initial sequencing and analysis of the human genome. *Nature 409:* 860-927.

Junqueira AC, Degrave W, Brandão A 2005. Minicircle organization and diversity in *Trypanosoma cruzi* populations. *Trends Parasitol 21*: 270-272.

Kadlec J, Izaurralde E, Cusack S 2004. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3 *Nat Struct Mol Biol 11*: 330-337.

Kazazian Jr HH 2000. L1 retrotransposons shape the mammalian genome. *Science 289*: 1152-1156.

Lauria-Pires L, Braga MS, Vexenat AC, Nitz N, Simões-Barbosa A, Tinoco DL, Teixeira ARL 2000. Progressive chronic Chagas heart disease ten years after treatment with anti-*Trypanosoma cruzi* nitroderivatives. *Am J Trop Med Hyg 63*: 111-118.

Liang P, Pardee AB 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science 257*: 967-971.

Liu B, Liu Y, Motyka SA, Agbo EE, Englund PT 2005. Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol 21*: 363-369.

Lukes J, Guilbride DL, Votypka J, Zikova A, Benne R, Englund PT 2002. Kinetoplast DNA network: evolution of an improbable structure. *Eukaryot Cell 1*: 495-502.

Lukka M, Tasa G, Ellonen P, Moilanen K, Vassiljev V, Ulmanen I 2005. Triallelic patterns in STR loci used for paternity analy-sis: Evidence for a duplication in chromosome 2 containing the TPOX STR locus. *Forensic Sci Internl* (Epub ahead of print).

Machado CA, Ayala FJ 2001. Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. *Proc Natl Acad Sci USA 98*: 7396-7401.

Marcus E 2005. Retraction controversy. *Cell 123*: 173-175.

Moser DR, Kirchhoff LV, Donelson JE. 1989. Detection of *Trypanosoma cruzi* by DNA amplification using the polymerase chain reaction. *J Clin Microbiol 27*: 1477-1482.

Mukherjee KK, Biswas R 2005. Short tandem repeat (STRs) and sex specific Amelogenin analysis of blood samples from neurosurgical female transfused patients. *J Clin Forensic Med 12*: 10-13.

Nitz N, Gomes C, Rosa AC, D'Souza-Ault MR, Moreno F, Lauria-Pires L, Nascimento RJ, Teixeira ARL 2004. Heritable integration of kDNA minicircle sequences from *Trypanosoma cruzi* into the avian genome: Insights into human Chagas Disease. *Cell 118*: 175-186.

Ostertag EM, Kazazian HH Jr 2001. Biology of mammalian L1 retrotransposons. *Ann Rev Genet 35*: 501-508.

Pavlícek A, Paces J, Elleder D, Hejnar J 2002a. Processed pseudogenes of human endogenous retroviruses generated by LINEs: Their integration, stability, and distribution. *Genome Res 12*: 391-399.

Pavlícek A, Paces J, Z1ka R, Hejnar J 2002b. Length distribution of long interspersed nuclear elements (LINEs) and processed pseudogenes of human endogenous retroviruses: implications for retrotransposition and pseudogene detection. *Gene 300*: 189-194.

Razin SV, Ioudinkova ES, Trifonov ED, Scherrer K 2001. Non-clonability correlates with genomic instability: A case study of a unique DNA region. *J Mol Biol 307*: 481-486.

Simonson AB, Servin JA, Skophammer RG, Herbold CW, Rivera MC, Lake JA 2005. Decoding the genomic tree of life. *Proc Natl Acd Sci USA 102*: 6608-6613.

Simpson AGB, Gill EE, Callahan HA, Litaker RW, Roger AJ 2004. Early evolution within kinetoplastids (Euglenozoa), and the late emergence of trypanosomatids. *Protist 155*: 407-422.

Sligtom JR, Bock JH, Tagle DA, Gumucio DL, Goodman M, Stojanovic N, Jackson J, Miller W, Hardison R 1997. The complete sequence of the galago and rabbit β-globin locus control regions: Extended sequence and functional conservation outside the cores of DNAse hypersensitive sites. *Genomics 39*: 90-94.

Smit AFA, Toth G, Riggs AD, Jurka J 1995. Ancestral mammalian wide subfamilies of LINE-1 repetitive sequences. *J Mol Biol 246*: 401-417.

Sturm NR, Degrave W, Morel C, Simpson L 1989. Sensitive detection and schizodeme classification of *Trypanosoma cruzi* cells by amplification of kinetoplast minicircle DNA sequences: use in diagnosis of Chagas disease. *Mol Biochem Parasitol 33*: 205-214.

Symer DE, Connely C, Szak ST, Capúto EM, Cost GJ, Parmigiani G, Boeke JD 2002. Human L1 retrotransposition is associated with genetic instability in vivo. *Cell 110*: 327-338.

Thertulien R, Simpson-Hidaris PG, Hidaris CG 1994. Intracellular localization of a *Trypanosoma cruzi* kDNA minicircle transcript using RNA:RNA in situ hybridization. *J Eukariotic Microbiol 41*: 402-407.

Teixeira ARL, Argañaraz ER, Freitas Jr LH, Lacava ZGM, Santana JM, Luna H 1994. Possible integration of *Trypanosoma cruzi* kDNA minicircles into the host cell genome by infection. *Mut Res 305*: 197-209.

Teixeira ARL, Nascimento RJ, Sturm NR 2006. Evolution and pathology in Chagas disease. *Mem Inst Oswaldo Cruz 101*: 463-491.

Venter CG, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA 2001. The sequence of the human genome. *Science 291*: 1304-1351.

Westenberger SJ, Barnabé C, Campbell DA, Sturm NR 2005. Two hybridization events define the population structure of *Trypanosoma cruzi. Genetics 171*: 527-543.

WHO-World Health Organization 1996. *The Global Burden of Disease,* In C Murray, A Lopez *(*eds), Geneva, Switzerland. Vol. I, chapter 5, p. 245-293.

WHO-World Health Organization 2002. Control of Chagas' disease: second report of a WHO Expert Committee. *Wld Hlth Org Tech Rep Ser 905*: 1-109.