Available online at www.sciencedirect.com
ScienceDirect
Molecular & Biochemical Parasitology 150 (2006) 236–248
MOLECULAR & BIOCHEMICAL PARASITOLOGY

Differential transcription profiles in *Trypanosoma cruzi* associated with clinical forms of Chagas disease: Maxicircle NADH dehydrogenase subunit 7 gene truncation in asymptomatic patient isolates☆

Cassio S. Baptista[a,1], Ricardo Z.N. Vêncio[a,b], Sarah Abdala[a], Julio César Carranza[a,c], Scott J. Westenberger[d], Marcelo N. Silva[a], Carlos A. de B. Pereira[b], Lúcia M.C. Galvão[e], Eliane D. Gontijo[f], Egler Chiari[e], Nancy R. Sturm[d], Bianca Zingales[a,*]

[a] *Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Avenida Professor Lineu Prestes, 748 CEP 05508-000, São Paulo, SP, Brazil*
[b] *BIOINFO-USP and Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, SP, Brazil*
[c] *LIPT, Universidad Del Tolima, Ibagué, Colombia*
[d] *Department of Microbiology, Immunology and Molecular Genetics, University of California at Los Angeles, Los Angeles, USA*
[e] *Departamento de Parasitologia, Instituto de Biociências, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*
[f] *Departamento de Medicina Preventiva, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*

Received 27 September 2005; received in revised form 18 August 2006; accepted 21 August 2006
Available online 15 September 2006

Abstract

The majority of individuals in the chronic phase of Chagas disease are asymptomatic (indeterminate form). Every year 2–3% of these individuals develop severe clinical manifestations (cardiac and digestive forms). In this study a *Trypanosoma cruzi* DNA microarray was used to compare the transcript profiles of six human isolates: three from asymptomatic and three from cardiac patients. Seven signals were expressed differentially between the two classes of isolates, including tryparedoxin, surface protease GP63, cyclophilin, some hypothetical proteins and the pre-edited maxicircle gene *NADH dehydrogenase subunit 7* (*ND7*). The approximately 30-fold greater signal in cardiac strains for *ND7* was the most pronounced of the group, and differential levels of pre-edited *ND7* transcript confirmed the microarray analysis. The *ND7* gene from asymptomatic isolates showed a deletion of 455 bp from nt 222 to nt 677 relative to *ND7* of the CL Brener reference strain. The *ND7* gene structure correlated with disease manifestation for 20 isolates from clinically characterised, chronic phase patients. The *ND7* lesion produces a truncated product that could impair the function of mitochondrial complex I. Possible links between the integrity of the electron transport chain and symptom presentation are discussed. We propose that *ND7* and other genes of the pathway constitute valuable targets for PCR assays in the differential diagnosis of the infective *T. cruzi* strain. While this hypothesis requires validation by the examination of additional recent parasite isolates from patients with defined pathologies, the identification of specific molecular markers represents a promising advance in the association between parasite genetics and disease pathology.

Keywords: DNA microarrays; Human strains; Diagnosis; NADH dehydrogenase subunit 7; Mitochondrial complex I; Tryparedoxin

1. Introduction

Chagas disease, caused by the protozoan *Trypanosoma cruzi*, affects 18 million people in Latin America. At present, a few hundred thousand seropositive individuals are estimated to live in the USA, Europe and Asia [1]. Chagas disease progresses through two successive stages: the acute phase and the chronic phase. The acute phase lasts 6–8 weeks and once it subsides most infected persons are asymptomatic. This presentation of the chronic phase is called the indeterminate form, and can persist indefinitely. In the majority of these individuals no abnormalities

Abbreviations: BER, Bayes error rate; COIII, cytochrome oxidase subunit III; ND7, NADH dehydrogenase subunit 7; EST, expression sequence tag; kDNA, kinetoplast DNA; LIT, liver infusion tryptose medium; TEUF and TENF, EST from a non-normalised (TEUF) and normalised (TENF) cDNA library of CL Brener epimastigotes

☆ *Note*: Nucleotide sequences reported in this paper have been submitted to the GenBank™ database with the accession numbers DQ663084–DQ663087.

* Corresponding author. Tel.: +55 11 3091 3810x217; fax: +55 11 3815 5579.
E-mail address: zingales@iq.usp.br (B. Zingales).
[1] Present address: Laboratory of Molecular Technology, 915 TollHouse, Ave. Suite 211, Frederick, MD 21702, USA.

0166-6851/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2006.08.008

are detected, while in others some functional alterations occur. The prognostic value of these alterations in the subsequent development of the chronic symptoms has not received appropriate attention. After several years or decades, 10–40% of the asymptomatic patients develop lesions in various organs, mainly the heart or digestive system. These presentations are called, respectively, the cardiac or digestive forms of Chagas disease [1].

The variability of symptoms and the geographical differences in the distribution of the chronic forms have been attributed to the diversity of *T. cruzi* strains. However the interplay with the immunogenetic background of the human host and the environment are also important for the outcome of Chagas disease. *T. cruzi* strains show substantial heterogeneity in biological and biochemical characteristics, which are resultant from a high genetic diversity (reviewed in [2]). Based on several DNA markers *T. cruzi* strains have been clustered into two major groups, which were named as *T. cruzi* I and *T. cruzi* II [3]. Further, *T. cruzi* II group has been divided into five subgroups IIa–IIe [4], several of which are the products of genetic fusions showing signs of recombination and heterozygosity [5]. Epidemiological studies suggest that *T. cruzi* I strains circulate in the silvatic cycle of the parasite transmission, whereas *T. cruzi* II strains are related to the domestic cycle in regions where Chagas disease is more severe [6].

In kinetoplastids the mitochondrial genome is represented by 20–50 maxicircles, which, together with thousands of minicircles, constitute a dense network, or kinetoplast (kDNA) that comprises approximately 15% of the total cellular DNA and thus represents an attractive diagnostic target. Genetic markers capable of differentiating *T. cruzi* II strains isolated from patients with cardiac and digestive forms were queried using a low-stringency single-primer PCR approach directed against the minicircle component of the kDNA, but it was not possible to identify a specific kDNA signature for strains isolated from a particular group of patients [7]. Minicircles carry the genes for a group of small transcripts referred to as guide RNAs that are involved in the Byzantine post-transcriptional process of RNA editing by uridine insertion and deletion [8]; due to the nature of their functional flexibility, these genes are largely transition tolerant, and may present hypervariability. The maxicircle genomes of the CL Brener and Esmeraldo strains have been assembled and annotated from data generated by the TIGR-SBRI-KI *T. cruzi* Sequencing Consortium [9]. Studies focusing on the phylogeny of the maxicircle relative to defined nuclear markers have defined three maxicircle clades [10,11] that are in agreement with an evolutionary schema proposed for *T. cruzi* including two major hybridisation events [5,9]. The functional constraints on maxicircle genes are also influenced by the RNA editing process, but these genes must maintain a higher degree of fidelity relative to the guide RNA genes.

One explanation for the distinct phenotypes and pathogenesis induced by *T. cruzi* strains may be the differential expression of particular genes. This hypothesis has been investigated in six strains classified into *T. cruzi* I and *T. cruzi* II groups by suppression subtractive hybridisation analysis [12]. Although diverse expression patterns were obtained for a few genes, no correlation between the gene expression and the classification of the strains was found. DNA microarray technology is a useful tool to discover new genes and to identify coordinated expression of genes. In a previous report, we demonstrated that despite the high genetic diversity of *T. cruzi* strains, DNA microarrays bearing predominantly expressed sequence tags (ESTs) of CL Brener are valid tools for comparative genomic studies and for the analysis of gene expression in this parasite [13]. In the present study a modified version of the microarray slide was constructed and used in the analysis of gene expression profiles in *T. cruzi* strains isolated from individuals presenting the indeterminate or the cardiac forms of Chagas disease. Among the DNA sequences differentially transcribed, the transcript from the mitochondrial maxicircle DNA-encoded gene for NADH dehydrogenase subunit 7 (*ND7*) was approximately 30-fold more abundant in the cardiac strains than in the asymptomatic strains. Detected differences were caused by a substantial deletion within *ND7*. The *ND7* locus was challenged by a set of *T. cruzi* isolates associated with characterised pathologies and maintained its validity. Thus the *ND7* gene may be a gateway for diagnostic and prognostic tests, and analysis of differentially expressed *T. cruzi* markers may advance the understanding of the basis of Chagas disease pathogenesis.

2. Materials and methods

2.1. Parasite stocks and cultivation

The characteristics of the *T. cruzi* strains employed in this study are summarised in Table 1. The strains were classified into *T. cruzi* I and *T. cruzi* II groups as previously described [14]. For the microarray experiments six Brazilian strains were analysed: the 115, B147, and B13-167 strains isolated from patients presenting severe electrocardiographic alterations [15], and the VL10, Famema, and Berenice 62 strains isolated from individuals with normal electrocardiograms and normal chest X-rays and considered to be asymptomatic (indeterminate form) (Table 1). The isolates were obtained by haemoculture in different years. In order to minimise parasite selection, positive blood cultures in liver infusion tryptose (LIT) medium with 10% fetal calf serum were maintained in individual tubes for a short period of time without passage. LIT medium was added every 10–15 days for a maximum of 8 weeks. Stocks were cryopreserved in liquid Nitrogen. For the validation of the PCR assay targeted at the *ND7* gene 25 isolates were analysed. The relevant information on reference strains was obtained from the literature (Table 1). These stocks were maintained in liquid Nitrogen (B.Z. laboratory) until use. Eleven human strains were isolated by haemoculture in the last 5 years from chronic patients from Minas Gerais (Brazil), who presented positive serology for Chagas disease (two tests). The clinical presentation of Chagas disease was meticulously assessed by one of us (E.D.G.) based on the results of electrocardiograms, echocardiograms, Holter monitoring, chest X-rays and radiology of the oesophagus and colon. This information will be published elsewhere. The patient characteristics are summarised in the lower section of Table 1. The same procedures described above were followed to minimise parasite selection. Pellets of 10^9 epimastigotes were frozen for DNA extraction.

Table 1
Characteristics of *T. cruzi* strains

Isolate	Group[a]	Origin	Host	Characteristics of the patient	Amplicon ND7 (bp)
115[b]	Tc II	Minas Gerais (BR)	Human; chronic phase; cardiac form	Adult, male	900
B147[b]	Tc II	Minas Gerais (BR)	Human; chronic phase; cardiac form	Adult, male	900
B13-167[b]	Tc II	Minas Gerais (BR)	Human; chronic phase; cardiac form	Adult, male	900
VL10[b]	Tc II	Minas Gerais (BR)	Human; chronic phase; indeterminate form	15-year old, female	500
Famema[b]	Tc II	São Paulo (BR)	Human; chronic phase; indeterminate form	57-year old, male	500
Berenice 62[b]	Tc II	Minas Gerais (BR)	Human; chronic phase; indeterminate form	53-year old, female	500
CL-Brener	Tc II	Rio Grande do Sul (BR)	*Triatoma infestans*	–	900
Hem 179	Tc II	Minas Gerais (BR)	Human; chronic phase; cardiac and digestive forms	60-year old, female	900
Colombiana	Tc I	Colombia	Human; chronic phase; cardiac form	NA	900
CA-1	Tc II	San Luis (AR)	Human; chronic phase; cardiac form	Adult, male	900
Silvio X10	Tc I	Pará (BR)	Human; acute phase	NA	900
Esmeraldo cl3	Tc II	Bahia (BR)	Human; acute phase	NA	900
Y	Tc II	São Paulo (BR)	Human; acute phase	NA	500
SC 2005	Tc II	Santa Catarina (BR)	Human; acute phase	Adult, male; recent oral contamination	900

Recent *T. cruzi* isolates from chronic chagasic patients from Minas Gerais (BR)[c]

Isolate	Clinical presentation	Characteristics of the patient	Amplicon ND7 (bp)
3 A	Indeterminate form	40-year old, female	500
81 J	Indeterminate form	40-year old, male	500
35 C	Indeterminate form	32-year old, female	500
45 D	Indeterminate form	40-year old, female	500
46 E	Indeterminate form	43-year old, female	500
42 F	Indeterminate form	45-year old, male	900 + 500
55 H	Indeterminate form	29-year old, female	900
54 G	Cardiac form	48-year old, male	900
54B	Cardiac form	48-year old, male	900
60 I	Cardiac and digestive forms	42-year old, male	900
83 K	Digestive (megacolon) and cardiac form (mild)	27-year old, female	500

NA, information not available.
[a] *T. cruzi* I and *T. cruzi* II major groups classified according to [14].
[b] Isolates used in the DNA microarray experiments.
[c] The clinical presentation of Chagas disease was characterised by several clinical exams (see Section 2). All isolates belong to *T. cruzi* II group.

2.2. Probes and construction of DNA microarrays

DNA microarrays were constructed with 710 ESTs obtained from non-normalised and normalised cDNA libraries of epimastigotes of the CL Brener strain [16] and 45 ESTs from a non-normalised cDNA library of amastigotes of the Tulahuen strain [17]. The array also contained 32 cloned *T. cruzi* RNA and protein coding genes and non-coding sequences from various strains, control sheared DNAs, and oligonucleotides [13]. Sequence similarity searches were performed with the BLAST version 2.2.6 suite of programs, against the non-redundant sequence database from NCBI (April 2006). Clustering of the 787 sequences with Cap3 program [18] revealed 57 clusters with 2–6 sequences each, 657 singletons and a total of 714 (90%) unique sequences. The probes were obtained as previously described [13] and immobilised on the glass slides with Generation III Microarray System spotter (Molecular Dynamics) according to manufacturer's instructions. The DNA probes were distributed in 10 sub-arrays, containing 288 spots each. Each sub-array was spotted in duplicate (left and right side of the slide) resulting in at least 6 and at most 60 replicates of each probe. After immobilisation, DNA was UV-irradiated at 254 nm (Stratagene lamp) with 50 mJ total energy. Complete information on the microarray slide can be found in GEO database [19] (http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE1828 with details required by MIAME guidelines [20].

2.3. Sample preparation, labelling and hybridisation

Total RNA was extracted with TRIzol reagent (Invitrogen) according to the manufacturer's protocol. For each strain, RNA was obtained from two independent parasite harvests of epimastigotes in mid-log growth phase. RNA was DNase-treated and approximately 20 μg was labelled with fluorophor-labelled dCTP in the first-strand cDNA synthesis. The reaction mix contained 0.05 mM of either Cy3- or Cy5-dCTP, 1 μl oligo-$dT_{(15)}$ primer (Amersham Biosciences), 4 μl of random nonamers (Amersham), 0.1 mM each of dATP, dCTP, dTTP and 0.05 mM dCTP, 10 mM dithiothreitol and 400 units of Superscript II Reverse Trascriptase (Invitrogen) in the buffer provided by the manufacturer. Incubation was performed at 42 °C for 2.5 h. Subsequently, RNA was hydrolysed by addition of 2 μl 2.5 M NaOH and an incubation at 37 °C for 15 min. The solution was neutralised by addition of 10 μl 2 M Hepes. The samples were purified with Multiscreen plates (Millipore). Incorporation measurement, hybridisation with the slides and washing conditions were as previously described [13]. Slides were dried and subjected to fluorescence detection.

2.4. Image acquisition and normalisation

The slides were scanned with a laser scanner (Molecular Dynamics), processed using ArrayVision 6.0 (Image Research) software, and the Cy3 and Cy5 intensities were re-parametrised as usual as the average log-intensity $A = \log 2(\text{Cy3})/2 + \log 2(\text{Cy5})/2$ and the log-ratios $M = \log 2(\text{probed strain/VL10})$, according to [13]. Low-intensity spots ($A \leq \text{mean} + 3$ standard deviation of negative controls) or saturated spots ($A \geq 9.5$) were excluded from the analyses. Normalisation of expression log-ratios M was achieved by LOWESS fitting as described in [13]. The complete experiment dataset, with raw intensity and normalised ratio data is publicly available in the GEO database [19] (http://www.ncbi.nlm.nih.gov/geo) by the accession number GSE1828.

2.5. Detection of differentially expressed genes

To identify differentially expressed signals between strains from asymptomatic and cardiac patients our premise was to select the genes with expression patterns that could statistically distinguish the three cardiac patient strains from the three asymptomatic patient strains. We defined as differentially expressed the probes that presented a Bayes error rate (BER) E value ≤0.05 [21,22]. This analysis was performed using the BayBoots software [23] with minor adaptations to generate outputs more specific to our dataset. Results for all the probes can be found at the supplementary web-site: www.vision.ime.usp.br/∼rvencio/Tcruzi/Cardiac.

2.6. RNA blotting

Total parasite RNA (10 μg) extracted with TRIzol reagent was separated on formaldehyde-containing 1% agarose gels in MOPS buffer according to standard procedures and blotted onto nitrocellulose membranes (Hybond-C, Amersham Pharmacia). DNA probes were labelled with [α-^{32}P] dATP with the Random Primer DNA Labelling Kit (Invitrogen). Hybridisation was carried out in 5× SSPE, 50% formamide, 5× Denhardt, 0.5% SDS at 42 °C, overnight. Blots were washed to final stringency of 0.1× SSPE, 0.1% SDS at 42 °C, and exposed to X-ray film. The radioactive images were also collected on Phosphor Screens (Kodak) and scanned with the Storm System (Molecular Dynamics). Densitometric analysis was performed with the ImageQuant Molecular Dynamics Program, and the data were normalised based on the hybridisation of the *T. cruzi* actin 2 probe. Prior to further probing, the blots were stripped by washing at 95 °C in 0.1× SSPE, 0.1% SDS. Membranes were exposed to Phosphor Screens to verify the efficiency of the process.

2.7. Southern blotting

Total parasite DNA was obtained as described [24]. kDNA was purified as described [25]. DNA preparations were digested with restriction enzymes (New England Biolabs), run on 0.8% agarose gels, and blotted to nylon membranes (Hybond-N, Amersham Pharmacia) using standard protocols. DNA probes were labelled as described above. Probes were hybridised at 60 °C in 0.1% Ficoll, 0.05% PVP, 1 mM EDTA, 3× SSC, 0.1% SDS and 100 μg ml^{-1} salmon sperm DNA. Blots were washed to a final stringency of 1× SSC, 0.1% SDS at 60 °C. The radioactive images were recorded on X-ray film or Phosphor Screens (Kodak). Prior to further probing, the blots were stripped by washing at 0.1× SSC, 0.5% SDS at 90 °C for 60 min.

2.8. PCR amplification of maxicircle sequences

Primers for amplification of *ND7* were designed in the genes flanking *ND7* based on the CL Brener sequence (GenBank accession no. DQ343645). The primers used for PCR amplification were Tc.Maxi.ND7.For (5′-AAGAAAAGAGGGGACAAACG-3′) and Tc.Maxi.ND7.Rev (5′-AAAAATCCCCTTCCAAAAGC-3′). PCR reaction conditions included initial 2 min denaturation at 93 °C, followed by 30 cycles at three temperatures: 30 s denaturation at 93 °C; 30 s primer annealing at 60 °C; and 1 min elongation at 72 °C; followed by 10 min elongation at 72 °C. PCR amplification of the *cytochrome oxidase subunit III* (*COIII*) sequence was performed with primers CO1 (5′-CGGAGCATTTTTATGGAGAGG-3′) and CO2 (5′-GAAGACCCTTCCTCTTTTCTCC-3′), designed to flank the *COIII* sequence of CL Brener (GenBank accession no. DQ343645). PCR reaction conditions included initial 4 min denaturation at 94 °C, followed by 35 cycles at three temperatures: 1 min denaturation at 94 °C; 1 min primer annealing at 57 °C; and 1 min elongation at 72 °C; followed by a 10 min elongation at 72 °C. PCR products were separated in 0.8% agarose gels and stained with ethidium bromide.

2.9. Sequencing of ND7 amplification products

The *ND7* PCR products were gel purified with Purelink Gel purification kit (Invitrogen). Purified PCR products were cloned using the TOPO-TA cloning kit (Invitrogen). Plasmid clones were purified with the Purelink PCR purification kit (Invitrogen) and sequenced commercially by Laragen (Los Angeles, CA) with plasmid-specific primers. Sequences were aligned using ClustalX and alignments were manually adjusted using BioEdit.

3. Results

3.1. DNA microarray assay with T. cruzi strains

To investigate differential gene expression between strains from cardiac and asymptomatic patients, we selected six strains typed as belonging to *T. cruzi* II group [3,14]. The three 'cardiac' strains (115, B147 and B13-167) were isolated from adult males of the same endemic region in Minas Gerais (Brazil) and who showed severe electrocardiographic alterations (Table 1). Two of the 'asymptomatic' strains (VL10 and Berenice 62) were from female patients of the same region as the three cardiac strains, and one (Famema) from a woman living in São Paulo State (Brazil). The Famema and Berenice 62 strains were from adults, while the VL10 was obtained from a 15-year old individual (Table 1). The asymptomatic patients showed normal electrocardiograms and chest X-rays.

To reduce the number of competitive hybridisations, the cDNA of the VL10 strain was arbitrarily used as reference. As such, six pairs of cDNA samples were hybridised with the microarray: VL10 × VL10; VL10 × 115; VL10 × B147; VL10 × B13-167; VL10 × Famema and VL10 × Berenice 62. Two sets of hybridisation experiments were performed with cDNA preparations obtained from two independent parasite harvests from the mid-log growth phase (biological replicates). At least six technical replicates of each probe were spotted in the microarray slide. Therefore, the results obtained for each probe refer to the hybridisation ratio R ($R = 2^{\bar{M}}$), where $\bar{M}$ is the mean of the M values of the two biological and all technical replicates in relation to the reference strain VL10.

3.2. Identification of differential gene expression

To identify genes that were differentially expressed between the two classes of strains, we looked for transcript signals that differed in strength and were common to the three cardiac strains when compared with the three asymptomatic strains. This identification was performed using the BayBoots method [23] that implements a model-free Bayesian statistical analysis. Briefly, this approach: (i) defines model-free density estimates of normalised ratios M of all the replicates from the asymptomatic class and from the cardiac class; (ii) calculates the Bayes error rate (BER) E value [21,22]; and (iii) sorts all probes by their E value. BER measures the distance of the M values for the replicates of each signal in the two classes of strains. E values of BER closer to zero reflect distant distributions, indicating consistent differential expression between the two classes. At the other extreme, E values closer to one correspond to superimposed distributions, indicating no differential behaviour between the two classes. In this work, the BER significance cutoff used was $E = 0.05$. Results for all the signals can be found at the supplementary web-site: www.vision.ime.usp.br/~rvencio/Tcruzi/Cardiac.

This method yielded 14 signals that were differentially expressed between the two classes of strains: 10 probes increased and 4 probes decreased in cardiac strains when compared to asymptomatic strains. The data of each probe are presented in Table 2 and refer to the hybridisation ratio ($R = 2^{\bar{M}}$, where $\bar{M}$ is the mean of the M values) in relation to the reference strain VL10 and the corresponding BER E value. The most striking result was a 17–30-fold increase of the signal of TENF0857 in cardiac strains. For the other signals lower hybridisation ratios (1.5–4.7) in relation to the VL10 strain were observed.

In Fig. 1 we show the M–A plots obtained for six sequences, four increased in cardiac strains (TENF0857; B6-Pl08_F10, TENF0426 and TENF0228) and two decreased in these strains (TENF0103 and TENF0853). The M–A plots indicate a homogeneous clustering of the biological and technical hybridisation replicates of the ESTs within the two classes of strains. Fig. 1 also shows the density estimates and the definition of the BER E values for the same ESTs (right insets). The profiles obtained indicate how distant the M hybridisation values of the replicates in the two classes of strains are. For TENF0857 and B6-Pl08_F10 a clear separation of the two profiles can be observed, with BER E values of zero. For the other ESTs the BER E values varied from 0.01 to 0.05 indicating that the separation, while significant, is less pronounced.

3.3. Direct RNA analysis of differentially expressed probes

To confirm the differential transcript expression, 11 of the 14 sequences shown in Table 2 were used as probes in blots of total RNA from the six human strains as well as RNA from the Hem 179 strain isolated from a patient with cardiac and digestive disorders (see Table 1). The first panel in Fig. 2 shows the migration of the 18S; 24Sα and 24Sβ rRNA molecules stained with ethidium bromide. The second panel shows the hybridisation of the blot with a probe of the *T. cruzi* house keeping gene actin 2. Quantification of the hybridisation signal by densitometric analysis using the ImageQuant Program indicated that the same amount of RNA was loaded for Famema, VL10, B147, B13-167 and Hem 179 strains, whereas 20% more RNA and 20% less RNA were applied, respectively, for the 115 and Berenice 62 strains.

Analysis of the hybridisation patterns shows that TENF0857 probe recognised a ~1 kb transcript only in cardiac strains (Fig. 2), confirming the data of the microarray experiments. In the case of the probes TENQ0823 and A4-Pl04_C6, the hybridisation profiles suggest higher abundance of these transcripts in cardiac strains. The inverse situation was observed for probe D2-Pl04_E01, which was more abundant in asymptomatic strains according to the microarray experiments. In this case, the probe hybridised with two transcripts in all the strains and the abundance of both mRNAs was increased in asymptomatic strains. For the probe B6-Pl08_F10, which was higher in cardiac strains

Table 2
Sequences differentially expressed in strains from cardiac and asymptomatic patients

Identity	GenBank	Similarity[a]	Hybridisation ratio in relation to VL10[b]						BER *E* value[c]	RNA blot[d]
			Asymptomatic strains			Cardiac strains				
			VL10	Famema	Berenice	115	B13-167	B147		
TENF0857	AI717801	kDNA, ND7	1.2	1.2	1.0	31.5	17.5	30.5	0	Confirmed
B6-Pl08_F10	CB923575	Aspartate aminotransferase	1.0	1.3	1.1	2.7	2.2	3.3	0	Not confirmed
TENF0305	AA532186	No match	ND	1.1	1.2	4.7	1.7	2.0	0	Confirmed
TENF0641	AA676105	Ribosomal protein L4	ND	1.2	1.0	4.0	1.5	1.5	0	Not confirmed
TENQ0823	CF243322	Hypothetical protein	1.0	1.1	1.2	1.9	1.4	1.6	0.02	Confirmed
A4-Pl04_C06	CB923896	tryparedoxin	1.3	1.1	1.3	1.7	1.7	2.1	0.02	Confirmed
TENF0593	AA676068	Surface protease GP63	1.1	1.1	1.0	2.3	1.5	1.8	0.03	Confirmed
TENF0426	AA532101	Hypothetical protein	1.0	1.0	1.1	1.5	1.2	1.8	0.04	–
TENU2842	AI069865	ribosomal protein S13	1.2	1.1	1.0	1.8	1.6	1.3	0.04	–
TENF0228	AA532138	Hypothetical protein	1.1	1.4	1.1	2.7	4.3	2.2	0.05	Confirmed
TENF0103	AA525713	Glucosamine-6-P isomerase	1.0	0.9	1.0	0.55	0.62	0.66	0.01	Not confirmed
TENF0853	AI717798	Hypothetical protein	0.9	1.0	0.9	0.6	0.71	0.71	0.03	Not confirmed
D2-Pl01_E01	CB923590	Cyclophilin	1.0	1.0	1.0	0.62	0.66	0.66	0.04	Confirmed
TENF0280	AA532175	No match	1.0	0.9	0.83	0.66	0.66	0.66	0.04	–

ND, not determined because the hybridisation replicates presented intensities below the cutoff (see Section 2).
[a] The best hit obtained by BLASTN and/or BLASTX programs against nr database of NCBI ($E \leq 10^{-5}$). All the matches were with *T. cruzi* genes/proteins.
[b] Hybridisation ratio ($R = 2^{\bar{M}}$) in relation to the reference strain VL10. $\bar{M}$ is the mean of the M values of all biological and technical replicates.
[c] Bayes error rate *E* value.
[d] Confirms or not the differential expression between the two classes of strains.

according to the microarray data, the RNA analysis did not confirm the result. In this case an interesting hybridisation pattern was observed: a 1.6-kb RNA was present in asymptomatic strains, whereas a 1.0-kb transcript was detected in the cardiac strains 115 and B147, and both transcripts of 1.6 and 1.0 kb were present in the B13-167 and Hem 179 strains. This pattern could correspond to the products of alleles of the same gene, cross-hybridisations with members of the same gene family, differential post-transcriptional processing, or maintenance of heterozygosity within some strains [5].

To better evaluate the abundance of the transcripts, the radioactive signal of the RNA blots hybridised with the 11 probes was quantified with the ImageQuant Program. When two or more transcripts were detected in a given strain, the areas of the hybridising bands were added. Data were normalised based on the hybridisation of the actin 2 probe. The RNA blot experiments confirmed differences in the abundance of transcripts in the two classes of strains for 7 of the 11 probes tested (summarised in Table 2). Sequence similarity searches with the BLAST programs against the non-redundant sequence database from NCBI showed that these probes correspond to *T. cruzi* genes coding for a chimeric fusion of NADH dehydrogenase subunit 7 (*ND7*) in maxicircle kDNA with a hypothetical protein in nuclear DNA, tryparedoxin, surface protease GP63, cyclophilin, and two hypothetical proteins with orthologs in the *T. brucei* and *L. major* genomes. One EST (TENF0305) showed no matches against *T. cruzi* coding sequences and was mapped to genomic contigs without annotation.

3.4. Genomic organisation of nuclear markers

In a previous study we examined whether the differential expression of a particular gene in two *T. cruzi* strains was related to differences in the gene copy number and/or sequence similarity between the strains [13]. Accordingly, we investigated the genomic organisation and abundance of five probes that had differential expression: tryparedoxin (A4-Pl04_C06); aspartate aminotransferase (B6-PI08_F10); hypothetical protein (TENQ0823); hypothetical protein (TENF0853); and cyclophilin (D2-Pl01_E01). Southern blots of total DNA digested with *Pst* I revealed no differences in the gene copy number between the strains (data not shown). These genes were not examined further in this study, but merit attention in the future.

3.5. Confirmation of the differential expression of ND7

The 330-bp EST TENF0857 (GenBank accession no. AI717801) showed the highest hybridisation ratios in cardiac strains in relation to the VL10 strain (ratios 17.5–31.5; see Table 2). RNA blots hybridised with this EST revealed the apparent absence of RNA transcripts in the asymptomatic strains. At the time we analysed this EST, the best BLASTN hits for the first 169 bp of TENF0857 were obtained with the kinetoplast maxicircle DNA of the Tulahuen strain (GenBank accession no. U43567; *E*-value 6×10^{-82}) and the kinetoplast *ND7* partial, pre-edited DNA sequence of the Silvio strain (GenBank accession no. U05881; *E*-value 3×10^{-68}). No significant similarity using BLASTX program was found. Positions 165–292 of EST TENF0857 showed the best BLAST hits with a nuclear hypothetical protein of CL Brener Tc00.1047053506401.140 (GenBank accession no. XM_812247; *E*-value 6×10^{-58}). This EST is an apparent cloning chimera of two genes, one nuclear and the other mitochondrial in location.

Because of the chimeric nature of EST TENF0857, we obtained a 325-bp EST (TEUF0045, GenBank accession no.

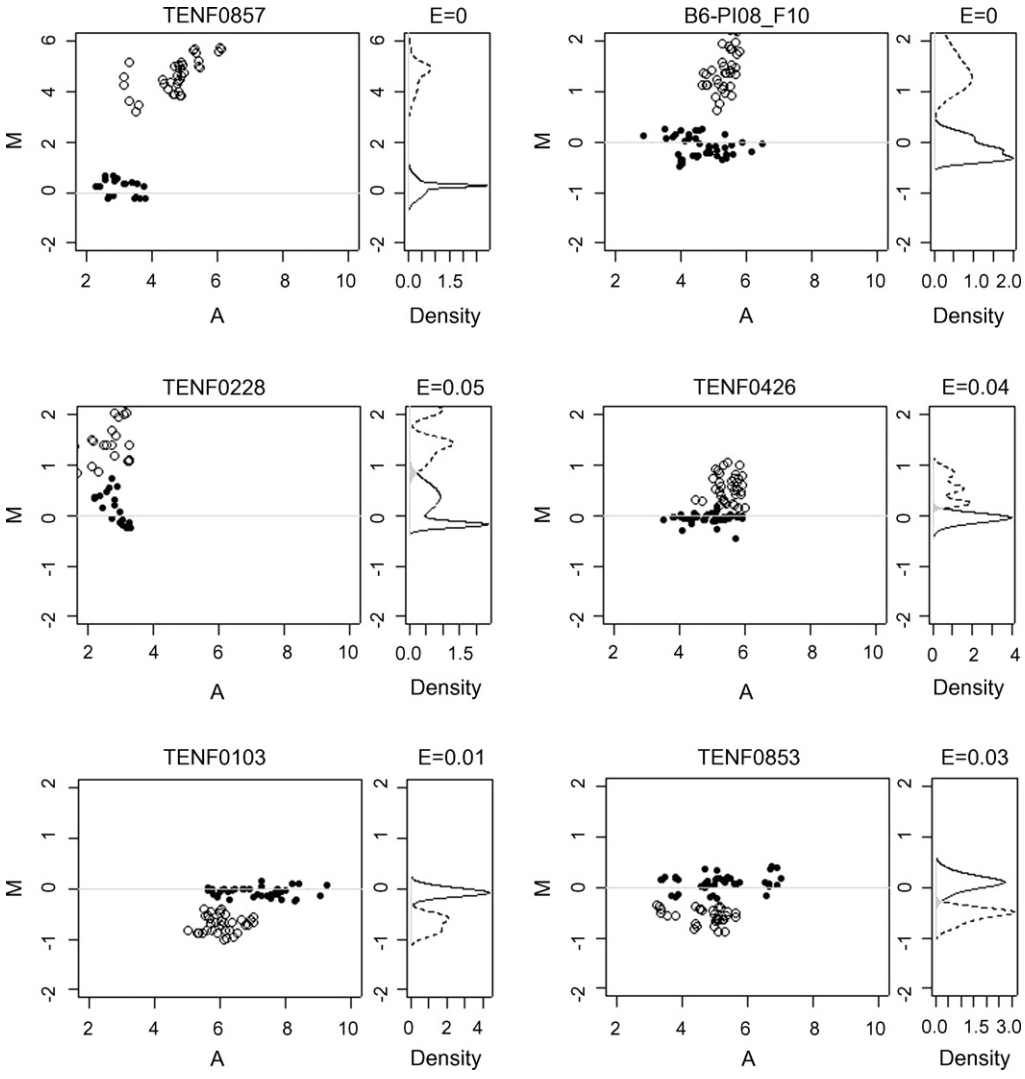

Fig. 1. Differential signal levels for chronic and asymptomatic associated *T. cruzi* strains by microarray analysis. *M* vs. *A* plots and determination of the Bayes error rate *E* value for six ESTs differentially expressed in three strains from cardiac patients (open circles and dotted lines) in relation to three strains of asymptomatic patients (solid circles and continuous lines). *M* is the hybridisation log-ratio ($M = \log 2(\text{probed strain}/\text{VL10})$); *A* is the average log-intensity ($A = \log 2(\text{Cy3})/2 + \log 2(\text{Cy5})/2$). The probability density is shown in the graphics on the right.

AA676164) that showed an improved BLASTN hit with the kinetoplast maxicircle DNA of the Tulahuen strain (GenBank accession no. U43567; *E*-value 3×10^{-99}) and *ND7* partial, pre-edited DNA sequence of the Silvio strain (GenBank accession no. U05881; *E*-value 6×10^{-73}), as observed for TENF0857. This EST is similar to ND7 over its entire length. To confirm the differential transcription of the *ND7* gene, the ND7 chimera TENF0857 and the pure ND7 TEUF0045 probes were hybridised to total RNA of the strains (Fig. 3). The hybridisation patterns confirmed the presence of a ~1 kb transcript only in the cardiac strains. Probe TEUF0045 gave a faint hybridisation of ~0.55 kb in the asymptomatic strains. This experiment also indicated that the gene encoding the hypothetical protein whose sequence is represented in the chimeric EST TENF0857 is not transcribed, as no transcript was detected in any of the isolates.

As a control for the expression levels of another maxicircle gene, we hybridised the same RNA blot with a probe against the adjacent maxicircle gene for cytochrome oxidase subunit III (*COIII*; TEUF0044; GenBank accession no. AA676163), an EST of CL Brener epimastigotes which presents similarity by BLASTN search with *T. cruzi* kinetoplast maxicircle DNA (Tulahuen strain) (GenBank accession no. U43567; *E* value 2×10^{-5}) and to *T. cruzi* Silvio X10 strain kinetoplast-encoded *COIII*, partial pre-RNA-edited DNA sequence (GenBank accession no. U05878; *E* value 3.5); the region of sequence identity begins at position 63 and extends through nucleotide 289 relative to the CL Brener maxicircle *COIII*. The original report of TEUF0044 in GenBank described similarity to the *T. cruzi* gene Histone H2B. In order to clarify this discrepancy, we sequenced approximately 600 bp of TEUF0044 and BLASTN analysis confirmed the assignment of pre-edited *COIII*. No significant similarity was found for TEUF0044 sequence in BLASTX searches against the NR database from NCBI. Therefore, TEUF0044 corresponds to *COIII*, which is contiguous to *ND7* in the maxicircle [9]. To verify whether *COIII* has a differential expression in the asymptomatic and cardiac strains, the RNA blot was hybridised

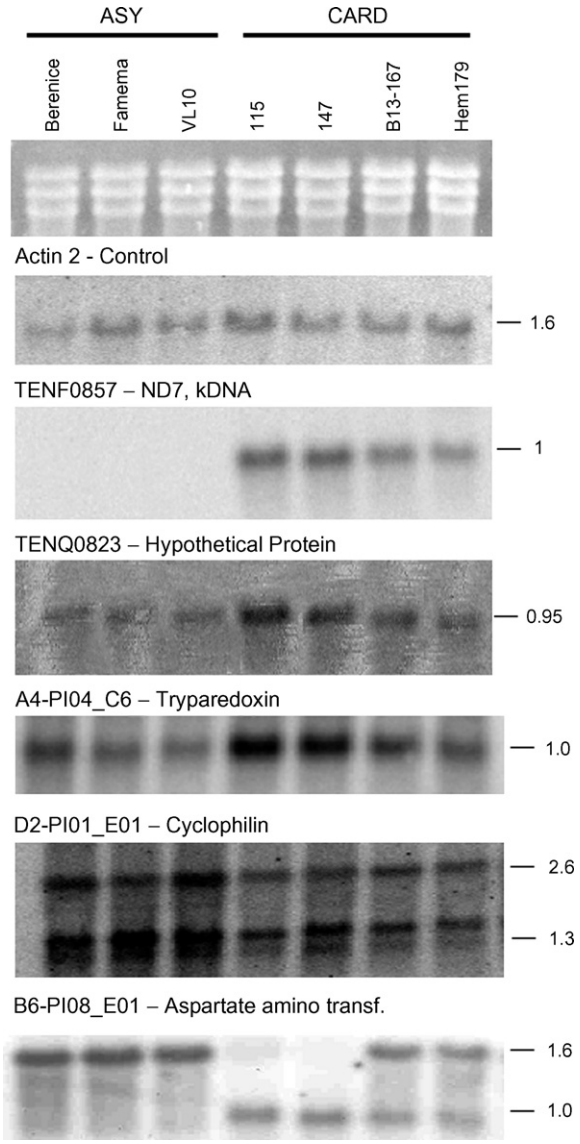

Fig. 2. Verification of microarray findings by direct RNA analysis. Blots of total RNA from *T. cruzi* strains. First panel: ethidium bromide staining of the region where the 18S; 24Sα and 24Sβ rRNAs migrate. Following panels: hybridisation with ^{32}P-labelled probes specified above each panel. The size of the transcript (in kb) is indicated on the right side.

to the ^{32}P labelled TEUF0044 probe (Fig. 3). A transcript of approximately 0.8 kb was observed in all the strains with equivalent intensities, indicating that *COIII* transcript accumulation was unaffected by the *ND7* situation.

3.6. The ND7 gene contains a deletion in asymptomatic strains

To investigate whether the differential expression of TENF0857 in cardiac and asymptomatic strains was due to differences in the abundance of *ND7*, we purified total DNA and kDNA from the VL10 (asymptomatic) and 115 (cardiac) strains as well as from the CL Brener strain. Total DNA (5 μg) and kDNA (1.7 μg) was digested with *Eco*RI and separated in an agarose gel (Fig. 4). The ethidium bromide staining indicated that similar amounts of DNA from the two sources were loaded. The bands of low molecular size visualised in the kDNA preparations correspond to minicircles, and showed distinct profiles between the cardiac and asymtomatic strains. After transfer to nylon membranes, the blot was hybridised to the *ND7* TEUF0045 probe. This probe identified a 6.8-kb band in the total DNA and kDNA of the 115 and CL Brener strains, and a band of 8.1 kb of lower intensity in the VL10 strain. This result indicated gross structural differences in *ND7* of the asymptomatic strain. The *COIII* probe hybridised with the same intensity with total DNA and kDNA of the three strains (Fig. 4), making it unlikely that maxicircles of VL10 are at a lower copy number. Differences in the genomic structure of *ND7* between the two classes of strains were confirmed by Southern blotting of DNA from two additional asymptomatic and two additional cardiac strains (data not shown). The combination of DNA and RNA analyses suggests that a deletion has occurred within *ND7* in the asymptomatic isolates, resulting in shorter hybridisation targets that yield reduced signals and variation in the size of genomic products (through loss of a restriction site) and the transcripts (due directly to the deletion) in the two populations.

3.7. Analysis of the ND7 deletion and sequencing

To examine differences in *ND7* structure among the strains, two PCR primers were designed in the genes flanking the *ND7* sequence from CL Brener [9] (GenBank accession no. DQ343645). Amplification of a ∼900 bp fragment of the expected size (897 bp) for *ND7* of CL Brener was produced using the DNA of the four cardiac strains; however, the three asymptomatic strains yielded a product of ∼500 bp (Fig. 5A). The *ND7* sequence was amplified in additional *T. cruzi* reference strains (Fig. 5B) typed in *T. cruzi* I and II groups (Table 1). In most of the strains a product of ∼900 bp was observed, whereas a ∼500 bp product was detected in the Y strain (Fig. 5B and upper section of Table 1). The *COIII* configuration of the cardiac and asymptomatic strains was also investigated by PCR amplification with primers designed to flank *COIII* of CL Brener [9] (GenBank accession no. DQ343645). A ∼700 bp fragment of the expected size (658 bp) was obtained with all strains examined (Fig. 5C).

The identical profiles of *ND7* from the asymptomatic strains, as well as the experimentally popular Y strain, was unexpected. In order to examine the molecular basis of the *ND7* deletion, the *ND7* PCR products from Silvio X10, 115, Y, and VL10 strains were cloned and sequenced. The sequences were aligned to the same region of the Esmeraldo and CL Brener maxicircle consensus sequences [9]. An identical deletion of 455 bp was detected in the VL10 and Y *ND7* sequences, spanning nt 222–nt 677 of the pre-edited *ND7* relative to CL Brener. The location of the deletion and the two EST probes is shown relative to the CL Brener sequence (Fig. 6).

The nature of the *ND7* deletion reveals why it was identified in the microarray analysis of differential transcript levels. The *ND7* sequence present on the chimeric probe lies entirely within the deletion common to the asymptomatic strains, result-

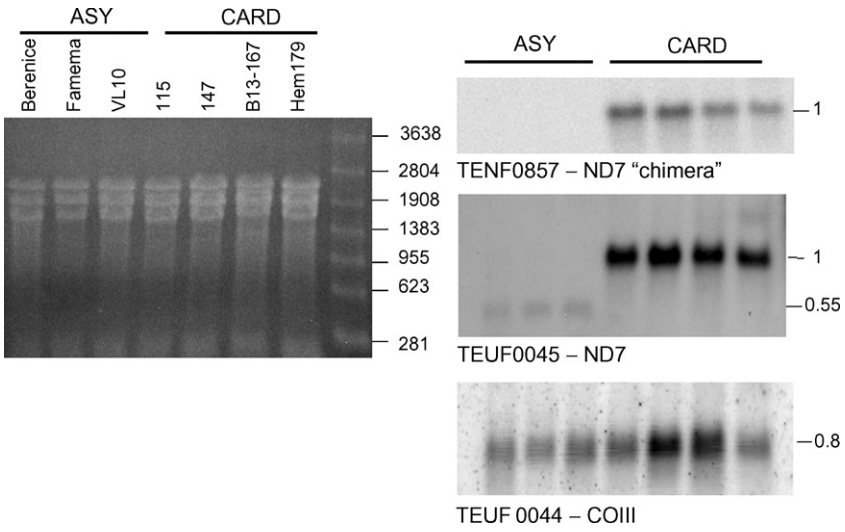

Fig. 3. Differential detection of maxicircle transcripts from *ND7* relative to *COIII*. Blot of total RNA from *T. cruzi* strains. Ethidium bromide staining of the region where the 18S; 24Sα and 24Sβ rRNAs migrate. Hybridisation with ^{32}P-labelled probes TENF0857, TEUF0045 and TEUF0044. Molecular mass markers are indicated.

ing in the stark contrast to the strains of cardiac origin. The other portion of the microarray probe did not detect transcript from either strain type, making its presence negligible. The relative placement of the second *ND7* EST overlapped the microarray *ND7* sequence in its entirety and extended beyond the 3′ end of the deletion in asymptomatic strains. This larger *ND7* probe was thus able to recognize the smaller transcripts produced in the asymptomatic strains as well as the full-length unedited transcript present in the cardiac strains (Fig. 3); the weak hybridisation of the asymptomatic transcripts can be related directly to the relatively small portion of the probe recognising the remaining sequence. Thus our microarray approach was successful in identifying a common difference amongst the two populations we set out to distinguish.

3.8. Validation of the ND7 deletion in asymptomatic strains

In order to expand the sample size beyond that used for the microarray study, additional clinical isolates were tested for anomalies in *ND7* by the straightforward amplification assay. Accordingly, we analysed the structure of *ND7* in 11 recent *T. cruzi* isolates from chronic patients from Minas Gerais (BR) (see details in Section 2). In these individuals the clinical presentation of Chagas disease was thoroughly characterised by multiple criteria.

The data shown in the upper section of Table 1 indicate that three isolates from asymptomatic patients had a deletion in *ND7*, while five isolates from cardiac patients and one isolate from a patient with mixed forms did not. The lower section of

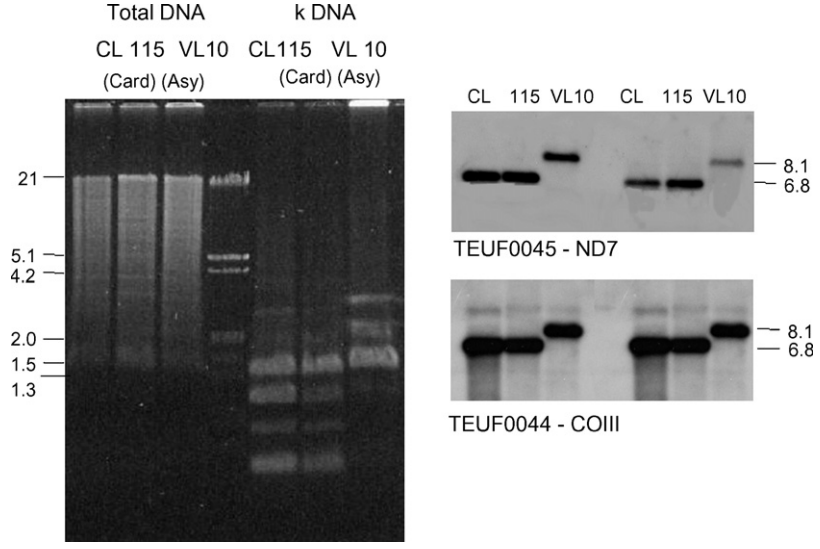

Fig. 4. Altered migration and reduced signal for maxicircle gene *ND7*. Southern blot of total DNA and kDNA of CL Brener (CL), 115 (cardiac) and VL10 (asymptomatic) strains digested with *Eco*RI. Ethidium bromide stained gel and hybridisations with ^{32}P-labelled probes TEUF0045 and TEUF0044. Molecular mass markers are indicated in kb.

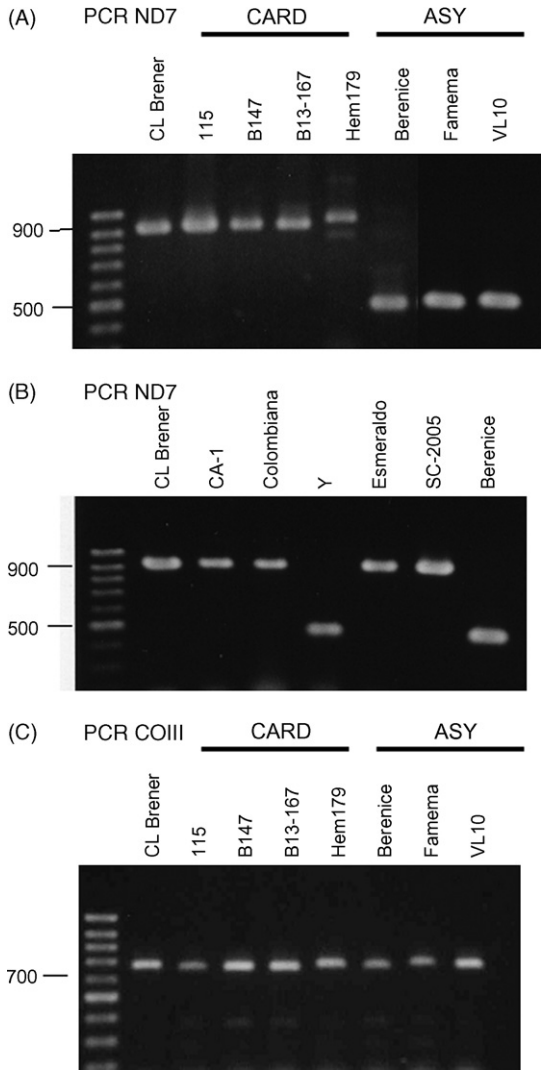

Fig. 5. A deletion in *ND7* gene associated with asymptomatic isolates. PCR amplification of the *ND7* (A and B) and *COIII* (C) sequences of *T. cruzi* strains. Ethidium bromide stained agarose gels. Molecular mass markers (in bp) are indicated.

Table 1 summarises the results found for the 11 fresh isolates. In five out of seven isolates from asymptomatic patients (normal electrocardiograms, echocardiograms, Holter and radiology of the oesophagus and colon), *ND7* amplicons of ∼500 bp were obtained. In two isolates from patients with cardiomyopathy and one isolate of a patient with cardiac and digestive manifestations, ∼900-bp amplicons were produced. The isolate from a patient with the digestive form (megacolon) and mild cardiac manifestations (normal echocardiogram and Holter; and right bundle branch block evidenced in the electrocardiogram) produced a ∼500-bp amplicon. The *T. cruzi* population from asymtpomatic patient '42 F' generated both the ∼500-bp and ∼900-bp amplicons, most probably reflecting the presence of at least two *T. cruzi* populations in this patient, and illustrating the discriminatory power of our assay. An apparently conflicting result was obtained for asymtpomatic isolate '55 H', which produced a ∼900-bp amplicon. This result does not invalidate our proposition that the deletion in *ND7* can be used for differentiation of the two classes of strains, but rather confirms the prognostic value of this test. According to our hypothesis, the individual who is presently asymptomatic and harbors a non-deleted *ND7 T. cruzi* strain is at risk of developing chagasic cardiomyopathy in the future, recommending periodic follow-up with the patient along with potential preventive treatment. An alternative, albeit less probable, possibility is that the deletion event in *ND7* may be a *de novo* occurrence in asymptomatic patients, or one of several molecular events leading to the same clinical result.

The data from these additional samples support the contention that the deletion in *ND7* constitutes a valuable target for PCR assays in the differential diagnosis of the infective *T. cruzi* strain and encourage further examination of the *ND7* locus from chronic patient isolates with solid clinical profiles.

4. Discussion

To identify transcripts differentially expressed between strains from cardiac and asymptomatic patients we employed microarray analysis. This approach identified 10 signals increased in cardiac strains and four signals increased in asymptomatic strains. For 13 probes of nuclear origin the hybridisation ratios ranged from 1.5 to 4.7 in relation to the VL10 reference strain. Moderate changes in mRNA abundance have been reported in most DNA microarrays studies in kinetoplastid parasites (*T. brucei*, *T. cruzi* and *Leishmania*) and have been attributed to the post-transcriptional control of gene expression in these organisms [26]. We confirmed differences in the abundance of mRNA between the two classes of strains for seven of the 11 tested signals. Differential transcript accumulation for five ESTs due to variation in gene copy number among the strains was ruled out. The potential for involvement in Chagas disease pathogenesis for the genes identified as differentially expressed at the transcript level including tryparedoxin, cyclophilin, putative surface protease GP63, and several hypothetical proteins, requires further directed studies.

Fig. 6. Schematic summary of the *ND7* region of the maxicircle. The CL Brener maxicircle numbering is used as a reference, with specific positions within *ND7* in parenthesis. The limit of the upstream region of RNA editing is symbolised by an anchor at nt 97 in the *ND7* transcript. The location of the deletion common to the asymptomatic strains is shown by a hatched box. Specific probes used in this study are shown above *ND7* and *COIII*; approximately placement of oligonucleotides used in PCR of ND7 are shown by lines with arrowheads. Regions of identity are represented by solid lines, while a dotted line shows the chimeric region of the microarray probe (see text).

The most impressive result obtained in the microarray experiments was the differential transcript accumulation from the mitochondrially localised *ND7* gene between the two classes of strains. Sequencing of this gene revealed a 455-bp deletion in the middle of the *ND7* coding region of the VL10 asymptomatic strain. This deletion results in a truncated mRNA and is responsible for the differential transcription observed in the microarray experiments. The *ND7* correlation was extended beyond the six strains used for the microarray analysis by the examination of additional *T. cruzi* strains isolated from clinically characterised patients. The follow-up of these patients as well as the analysis of additional *T. cruzi* isolates from adult asymptomatic and symptomatic patients will further challenge our hypothesis.

In this study it calls the attention the size dimorphism of the *ND7* amplicon in *T. cruzi* isolates and conservation of the deletion in the *ND7* amplicon, demonstrated by sequencing of the VL10 and Y strains and suggested by the consistent sizes of deleted products. The conservation of the mutation in the *ND7* maxicircle gene proposes this is an ancient event and that the isolates with the deletion descend from a common ancestor wherein this mutation arose; alternatively, common sequence elements flanking the deletion may favor the truncated *ND7* form here described. Future verification of the deletions occurring in the other strains carrying the ~500 bp product, and characterisation of genetic markers shared by this putative sub-group of strains must be performed to evaluate any evolutionary implications. The isolates with deletion in *ND7* belong to the broad *T. cruzi* II group, and further genotyping has not yet been performed. Two reference strains from DTU IIb (Y and Esmeraldo cl3) produced different amplicons, suggesting that *ND7* may discriminate within the DTU IIb group similar to the *5S rRNA* array [27], or that the deletion can arise spontaneously in any background.

The 455-bp deletion we report in the middle of *ND7* from asymptomatic strains results in a truncated mRNA. This deletion cannot be overcome by RNA editing, nor was evidence of residual full-length *ND7* copies detected. In fact, the *ND7* transcript requires extensive RNA editing in *T. brucei* [28], with two large regions of editing separated by an unedited region of approximately 70 bp. The precise pattern of editing in *T. cruzi* has not yet been determined, however the unedited 70-bp block is conserved between *Trypanosoma* and *Leishmania*. This unedited region provides the anchor sequence through which the guide RNA recognises the mRNA. As the 5′-anchor region lies upstream of the deletion in the asymptomatic strains, the editing that initiates from that point may yet occur, potentially creating a start codon for the truncated transcript.

ND7 encodes the pre-edited backbone of the ND7 subunit of respiratory complex I, which couples electron transfer from NADH to ubiquinone. Complex I is a large multimolecular entity containing more than 40 subunits in mammals and plants [29]. In plants the NAD7 and NAD9 subunits are probably located in the matrix arm and the NAD7 subunit has been suggested to participate directly in NADH oxidation. Comparison of the N-terminal amino acid sequence of the NAD7 subunit from potato [30] and the sequence deduced from the maxicircle *ND7* of trypanosomes [28] suggests that the ND7 subunit of *T. cruzi* would be involved directly in the oxidation of NADH. In face of the deletion observed in *ND7*, some dysfunction of complex I can be predicted. On the other hand, *T. cruzi* strains with deletions in *ND7* display normal growth in LIT medium and infect mammalian hosts. This observation suggests that complex I alterations are tolerated in *T. cruzi* strains. The presence of a complex I-like NADH dehydrogenase in *T. cruzi* and trypanosomatids has long been debated (reviewed in [31]). The presumed absence of complex I activity was based predominantly on the observed absence of rotenone-sensitive NADH oxidation. However, not all complex I incarnations have the same high sensitivity towards this inhibitor as the mammalian type [32]. Therefore, the determination of complex I activity in *T. cruzi* strains with alterations in the genes encoding complex I subunits will contribute to the elucidation of metabolic aspects of this parasite.

In mammals, complexes I and III are the main sites of mitochondrial superoxide production [33,34]. In *T. cruzi* mitochondrial respiration generates superoxide [35], which is also formed by redox cycling of antichagasic drugs [36]. Future studies will determine which of the two classes of strains produces higher levels of mitochondrial reactive oxygen species, and which is more efficient in detoxification. Notably, tryparedoxin, a component of the trypanothione peroxidase pathway whose principal biological function is to control oxidative stress ([37] and references therein), is 1.7–2.0-fold up regulated in cardiac strains. Peroxynitrite is a strong oxidizing and cytotoxic effector molecule against *T. cruzi* [38] and has been reported in the myocardium of acute chagasic experimental animals [39,40]. Furthermore, profiling of gene transcription in *T. cruzi*-infected murine hearts has shown repression of host genes encoding various components of the electron transport chain [41]. Dysfunction of the oxidative phosphorylation pathway may be responsible for production of reactive oxygen species in the myocardium. The increase of tryparedoxin in cardiac strains could explain to some extent why reactive oxygen species are not sufficient to clear the parasites in the myocardium, thus contributing to the pathogenesis of chagasic cardiomyopathy.

The molecular mechanisms of pathogenesis underlying the chronic phase of Chagas disease are poorly understood. Progression to the chronic disease is very likely a multifactorial process that depends on the combination of host genetic factors, parasite genetic factors, and environmental factors. Few studies have compared specific aspects of the immune response in asymptomatic patients and patients with chronic Chagas cardiomyopathy (reviewed in [42]). No association has been verified between the risk of developing cardiomyopathy and two HLA class II genes in the Brazilian population [43]. More recently, two polymorphisms of the *BAT1* gene (a member of the DEAD-box family of RNA helicases) have been associated with Chagas cardiomyopathy [42]. These polymorphisms are either functionally associated with the development of cardiomyopathy and, thus, reflect true association, or are in linkage disequilibrium with a critical gene.

In this study we have shown that the *ND7* gene is an interesting genetic marker to differentiate cardiac and asymptomatic strains. A prognostic test would be valuable for the follow-up of the clinical manifestation, dictating adequate medical care.

We cannot associate directly the putative alteration of complex I activity with pathogenesis, although we suggest that *ND7* is a marker for a potentially interesting symptom that could manifest itself in conjunction with the expression of other genes, rather than a definitive marker in itself.

Acknowledgements

We thank Dr. Wim Degrave, Dr. Adeilton Brandão and Dr. Santuza Teixeira for the EST clones and Adriana Matsukuma for technical assistance. We are indebted to Dr. Daniella Bartholomeu, Dr. Gustavo Cerqueira, Dr. Nagib El-Sayed, and Dr. David Campbell for insightful suggestions. This work was supported by grants of Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (BZ); PRONEX-FAPEMIG/CNPq (EC) and the UCLA Bioinformatics Integrative Graduate Education and Research Traineeship program funded by NSF grant DGE9987641 (SW) and NIH award AI056034 (NS).

References

[1] WHO Expert Committee on the Control of Chagas Disease. Second report. Geneva: World Health Organization;2002.
[2] Campbell DA, Westenberger SJ, Sturm NR. The determinants of Chagas disease: connecting parasite and host genetics. Curr Mol Med 2004;4:549–62.
[3] Anonymous recommendations from a satellite meeting. Mem Inst Oswaldo Cruz 1999;94:429–32.
[4] Brisse S, Barnabé C, Tibayrenc M. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. Int J Parasitol 2000;30:35–44.
[5] Westenberger SJ, Barnabé C, Campbell DA, Sturm NR. Two hybridization events define the population structure of *Trypanosoma cruzi*. Genetics 2005;171:527–43.
[6] Zingales B, Souto RP, Mangia RH, et al. American Trypanosomiasis in Brazil based on dimorphism of rRNA and mini-exon sequences. Int J Parasitol 1998;28:105–12.
[7] Vago AR, Andrade LO, Leite AA, et al. Genetic characterization of *Trypanosoma cruzi* directly from tissues of patients with chronic Chagas disease. Am J Pathol 2000;156:1805–9.
[8] Simpson L, Sbicego S, Aphasizhev R. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. RNA 2003;9:265–76.
[9] Westenberger SJ, Cerqueira GC, El-Sayed NM, Zingales B, Campbell DA, Sturm NR. *Trypanosoma cruzi* mitochondrial maxicircles display species- and strain-specific variation and possess a conserved element in the non-coding region. BMC Genomics 2006;7:60.
[10] Machado CA, Ayala FJ. Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. Proc Natl Acad Sci USA 2001;98:7396–401.
[11] Brisse S, Henriksson J, Barnabé C, et al. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. Infect Genet Evol 2003;2:173–83.
[12] Dost CK, Saiva J, Monesi N, Zentgraf U, Engels W, Albuquerque S. SIX *Trypanosoma cruzi* strains characterized by specific gene expression patterns. Parasitol Res 2004;94:134–40.
[13] Baptista CS, Vêncio RZN, Abdala S, et al. DNA microarrays for comparative genomics and analysis of gene expression in *Trypanosoma cruzi*. Mol Biochem Parasitol 2004;138:183–94.
[14] Souto RP, Fernandes O, Macedo AM, Campbell DA, Zingales B. DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. Mol Biochem Parasitol 1996;83:141–52.
[15] Zingales B, Abuin G, Romanha AJ, Chiari E, Colli W. Surface antigens of stocks and clones of *Trypanosoma cruzi* isolated from humans. Acta Trop 1984;41:5–16.
[16] Ürmenyi TP, Bonaldo MF, Soares MB, Rondinelli E. Construction of a normalized cDNA library for the *Trypanosoma cruzi* Genome Project. J Eukaryot Microbiol 1999;46:542–4.
[17] Darocha WD, Bartholomeu DC, Macêdo CDS, et al. Characterization of cDNA clones encoding ribonucleoprotein antigens expressed in *Trypanosoma cruzi* amastigotes. Parasitol Res 2002;88:292–300.
[18] Huang X, Madan A. CAP3: a DNA sequence assembly program. Genome Res 1999;9:868–77.
[19] Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridisation array data repository. Nucleic Acids Res 2002;30:207–10.
[20] Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. Nat Genet 2001;29:365–71.
[21] Duda RO, Hart PE, Stork DG. Pattern classification. 2nd ed. Wiley/Interscience; 2000.
[22] Vêncio RZN, Brentani H, Patrão DFC, Pereira CAB. Bayesian model accounting for within-class biological variability in SAGE analysis. BMC Bioinform 2004;5:119.
[23] Vêncio RZN, Patrão D, Baptista CS, Pereira CA de B, Zingales B. BayBoots: a model-free Bayesian tool to identify class markers from gene expression data. Genet Mol Res 2006;5:138–42.
[24] Medina-Acosta G, Cross GAM. Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. Mol Biochem Parasitol 1993;59:327–30.
[25] Gonçalves AM, Nehme NS, Morel CM. An improved silver staining procedure for schizodeme analysis in polyacrylamide gradient gels. Mem Inst Oswaldo Cruz 1990;85:101–6.
[26] Duncan R. DNA microarray analysis of protozoan parasite gene expression: outcomes correlate with mechanisms of regulation. TRENDS Parasitol 2004;5:211–5.
[27] Westenberger SJ, Sturm NR, Campbell DA. *Trypanosoma cruzi 5S rRNA* arrays define five groups and indicate the geographic origins of an ancestor of the heterozygous hybrids. Int J Parasitol 2006;36:337–46.
[28] Koslowsky DJ, Bhat GJ, Perrollaz AL, Feagin JE, Stuart K. The MURF3 gene of *T. brucei* contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. Cell 1990;62:901–11.
[29] Heazlewood JL, Howell KA, Millar AH. Mitochondrial complex I from *Arabidopsis* and rice: orthologs of mammalian and fungal components coupled with plant-specific subunits. Biochim Biophys Acta 2003;1604:159–69.
[30] Gabler L, Herz U, Liddell A, et al. The 42.5 kDa subunit of the NADH: ubiquinone oxidoreductase (complex I) in higher plants is encoded by the mitochondrial nad7 gene. Mol Gen Genet 1994;244:33–40.
[31] Cazzulo JJ. Intermediate metabolism in *Trypanosoma cruzi*. J Bioenerg Biomembr 1994;26:157–64.
[32] Tielens AG, Van Hellemond. Reply. Parasitol Today 1999;15:347–8.
[33] Votyakova TV, Reynolds IJ. DeltaPsi(m)-dependent and -independent production of reactive oxygen species by rat brain mitochondria. J Neurochem 2001;79:266–77.
[34] Muller FL, Liu Y, Remmen VL. Complex III releases superoxide to both sides of the inner mitochondrial membrane. J Biol Chem 2004;47:49064–73.
[35] Boveris A, Stoppani AO. Hydrogen peroxide generation in *Trypanosoma cruzi*. Experientia 1977;33:1306–8.
[36] Docampo R, Moreno SN. Free radical metabolism of antiparasitic agents. Fed Proc 1986;45:2471–6.
[37] Alphey MS, Gabrielsen M, Micossi E, et al. Tryparedoxins from *Crithidia fasciculata* and *Trypanosoma brucei*. J Biol Chem 2003;278:25919–25.
[38] Thompson L, Denicola A, Radi R. The trypanothione-thiol system in *Trypanosoma cruzi* as a key antioxidant mechanism against peroxynitrite-mediated cytotoxicity. Arch Biochem Biophys 2003;412:55–64.

[39] Chandrasekar B, Melby PC, Troyer DA, Freeman GL. Differential regulation of nitric oxide synthase isoforms in experimental acute chagasic cardiomyopathy. Clin Exp Immunol 2000;121:112–9.

[40] Machado FS, Martins GA, Aliberti JC, Mestriner FL, Cunha FQ, Silva JS. *Trypanosoma cruzi*-infected cardiomyocytes produce chemochines and cytokines that trigger potent nitric oxide-dependent trypanocydal activity. Circulation 2000;102:3003–8.

[41] Garg N, Popov VL, Papaconstantinou J. Profiling gene transcription reveals a deficiency in mitochondrial oxidative phosphorylation in *Trypanosoma cruzi*-infected murine hearts: implications in chagasic myocarditis development. Biochim Biophys Acta 2003;1638:106–20.

[42] Ramasawmy R, Cunha-Neto E, Fae KC, et al. BAT1, a putative anti-inflammatory gene, is associated with chronic Chagas cardiomyopathy. J Infect Dis 2006;193:1394–9.

[43] Fae KC, Drigo SA, Cunha-Neto E, et al. HLA and beta-myosin heavy chain do not influence susceptibility to Chagas disease cardiomyopathy. Microbes Infect 2000;2:745–51.