Chromosomal size variation in *Trypanosoma cruzi* is mainly progressive and is evolutionarily informative

J. HENRIKSSON[1], J. C. DUJARDIN[2]*, C. BARNABÉ[3], S. BRISSE[4], G. TIMPERMAN[2], J. VENEGAS[5], U. PETTERSSON[1], M. TIBAYRENC[3] *and* A. SOLARI[5]

[1] *Department of Medical Genetics, Uppsala University, Biomedical Center, Box 589, S-751 23 Uppsala, Sweden*
[2] *Laboratory of Protozoology, Institute of Tropical Medicine 'Prince Leopold', Nationalestraat 155, B-2000 Antwerpen 1, Belgium*
[3] *Génétique des Maladies Infectieuses, UMR CNRS/IRD 9926, IRD, BP5045, F-34032 Montpellier Cedex, France*
[4] *Eijkman-Winkler Institute, University Medical Centre Utrecht, AZU G04.614, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*
[5] *Program of Cellular and Molecular Biology, Institute of Biomedical Sciences, Faculty of Medicine, University of Chile, Casilla 70086, Santiago-7, Chile*

(*Received 3 July 2001; revised 25 September 2001; accepted 26 September 2001*)

SUMMARY

The evolutionary significance of chromosome size polymorphism was explored in a representative panel of 26 *Trypanosoma cruzi* stocks. We tested a progressive model (aCSDI) assuming that the larger the size difference between homologous chromosomes, the more divergent the parasites are. This was contrasted with a non-progressive model (Jaccard's distance), in which any chromosome size difference has the same weight. ACSDI-based dendrograms were very similar to those built-up from multilocus enzyme electrophoresis (MLEE) and random amplified polymorphic DNA (RAPD) data: structuring in 2 major lineages (*T. cruzi* I and *T. cruzi* II) and 5 small subdivisions within *T. cruzi* II was identical, and branching was very similar. Furthermore, a significant correlation ($P < 0{\cdot}001$) was observed between aCSDI and phenetic distances calculated from MLEE and RAPD data. In contrast, analysis of chromosome size polymorphism with Jaccard's distance generated dendrograms with relatively long branches, causing most branching points to cluster close together, which generates statistically uncertain branching points. Our results thus support a model of progressive chromosome size-variation and show that despite an extensive polymorphism, chromosomal sizes constitute valuable characters for evolutionary analyses. Furthermore, our data are consistent with the clonal evolution model previously proposed for *T. cruzi*.

Key words: molecular karyotype, isoenzymes, RAPD, phenetic analysis, *Trypanosoma cruzi*.

INTRODUCTION

Trypanosoma cruzi is the causative agent of Chagas disease, which is one of the major parasitic diseases affecting humans in South America. *T. cruzi* has a broad host and vector range; it can infect a large number of different mammal species and it is transmitted by several species of blood-sucking reduviid bugs. The disease has been estimated to affect 16–18 million people in South America (World Bank, 1993), but control programmes are in progress (Schofield & Dias, 1999). The clinical manifestations of Chagas disease show considerable variation in affected individuals, from asymptomatic to severe symptoms including death. Interestingly, there is correlation between geographical location and the clinical symptoms of the disease (Miles *et al.* 1981; Brener, 1982). This geographical correlation might depend on several different factors, e.g. differences between parasite strains (Laurent *et al.* 1997), insect vector and human populations in different areas.

An extensive heterogeneity among different *T. cruzi* isolates has been described using Multilocus Enzyme Electrophoresis (MLEE) and DNA analyses, leading in the past to different attempts of classification. Last recommended nomenclature (Momen, 1999) refers to 2 major lineages, *T. cruzi* I and *T. cruzi* II. Furthermore, more recent studies by MLEE, RAPD, 24Sα rRNA, 18S rRNA and mini-exon PCR markers allowed the unambiguous identification of 5 smaller subdivisions within *T. cruzi* II (*T. cruzi* IIa–e; Barnabé, Brisse & Tibayrenc, 2000; Brisse, Barnabé & Tibayrenc, 2000; Brisse, Verhoef & Tibayrenc, 2001). The strong genetic structuring observed within *T. cruzi* is parsimoniously explained by long-term clonal evolution with only rare events of genetic hybridization (Tibayrenc, 1995).

Heterogeneity of *T. cruzi* was also observed at the level of molecular karyotyping (Engman *et al.* 1987; Henriksson *et al.* 1990). In a previous study including 46 *T. cruzi* stocks (Henriksson, Petterson & Solari, 1993), samples with similar MLEE profiles showed identical or a very similar molecular karyo-

* Present and corresponding address: Laboratory of Protozoology, Institute of Tropical Medicine 'Prince Leopold', Nationalestraat 155, B-2000 Antwerpen 1, Belgium. Tel: +32 3 2476355. Fax: +32 3 2476359. E-mail: jcdujard@itg.be

 © 2002 Cambridge University Press
DOI: 10.1017/S0031182001001093 Printed in the United Kingdom

type. Reciprocally, samples with different MLEE profiles showed a different molecular karyotype. However, the evolutionary significance of karyotype variation was never explored, mostly due to the absence of an underlying theoretical model.

A first insight in karyotype evolution can be gained by analysis of chromosome size polymorphism. In Trypanosomatids, this phenomenon is frequent and mostly due to expansion/contraction of tandem repeats (Wagner & So, 1990; Campetella *et al*. 1992; Åslund *et al*. 1994; Henriksson *et al*. 1995, 1996; Victoir *et al*. 1995; Inga *et al*. 1998; Kebede *et al*. 1999). Chromosome fusion/fission events might also play a role, albeit occurring very rarely (Britto *et al*. 1998). Accordingly, we assumed that the larger the size difference between homologous chromosomes from different stocks, the more divergent the stocks are (progressive model). On this basis, we designed the absolute chromosomal size difference index (aCSDI), a phenetic method allowing to weight size differences (Dujardin *et al*. 1995). This method contrasts with presence/absence approaches like the distance of Jaccard (1908), in which any size difference has the same weight (non-progressive model). ACSDI was previously applied to *Leishmania* populations and shown to be relevant for building evolutionary hypotheses at intraspecific (Dujardin *et al*. 1995, 1998) and genus (Dujardin *et al*. 2000) levels.

In the present study, we aimed to (i) analyse the evolutionary significance of chromosome size-variation in *T. cruzi*, and (ii) determine if the progressive model is more appropriate to represent the process of chromosome evolution in *T. cruzi*. Therefore, we applied the aCSDI model to the analysis of 7 chromosomes in a representative panel of 26 *T. cruzi* cloned stocks. Results were compared with those obtained by MLEE and RAPD on one hand, and by Jaccard's analysis of chromosomal size on the other hand.

MATERIALS AND METHODS

Parasites

Epimastigote forms of *T. cruzi* were grown on liquid medium supplemented with 5% fetal calf serum (Diamond, 1968). Information about the samples listed in Table 1 is taken from reports by Tibayrenc & Ayala (1988) or Tibayrenc *et al*. (1993) with the exception for Sylvia X10 cl.7 (Dvorak *et al*. 1982), Sp104 cl.1 (Apt *et al*. 1987; Solari *et al*. 1992), and CL Brener (Goldberg & Silva Pereira, 1983). Throughout the present work, we adopted the recommended nomenclature (Momen, 1999) for *T. cruzi* groups, i.e. *T. cruzi* I and II. However, this classification was completed by subdivision of *T. cruzi* II into 5 smaller subdivisions (IIa–e) according to Barnabe *et al*. (2000) and Brisse *et al*. (2000, 2001). For comparison with previous works, we included in Table 1 the equivalence to classification by Miles *et al*. (1980; zymodemes I, II and III) and Tibayrenc *et al*. (1986, clonets 1–43).

DNA preparation and electrophoresis

Agarose-embedded DNA samples for PFGE were prepared as described by Engman *et al*. (1987) and in each gel slot DNA corresponding to approximately $5–10 \times 10^6$ epimastigotes was loaded. All chromosomal separations were performed in a Chef Mapper apparatus (Bio-Rad), using the conditions described for PFGE separating programme I (Henriksson *et al*. 1995) which generates reproducible separations of good quality with an almost linear size separation of DNA molecules of sizes between 400 and 1800 kb. As size markers *Saccharomyces cerevisiae* and *Hansenula wingei* chromosomes (Bio-Rad) were used. After the separation had been completed, the DNA was transferred to nylon filters (Pall Biodyne) as described (Henriksson *et al*. 1990).

Probes and hybridization

Seven chromosomes, shown to constitute conserved linkage groups (Henriksson *et al*. 1995) were hybridized with specific markers: 4 gene fragments, 1F8 (Gonzalez *et al*. 1985), cruzipain (Åslund *et al*. 1991), FFAg6 (Porcel *et al*. unpublished observations) and Tc2 (Ibañez *et al*. 1988), and 3 anonymous genomic markers, CA7.12, CA7.32 and P19 (Henriksson *et al*. 1995). The selected markers hybridized to small or intermediate-sized chromosomes ranging between 400 and 1800 kb in all isolates, with only 2 exceptions (FFAg6 marker hybridized to large chromosomes migrating in the compression zone of the gel in the b2148 and the b2149 stocks). The probes were ^{32}P labelled using the Rediprime kit (Amersham). Nylon filters containing separated chromosomes were hybridized and washed at high stringency conditions as described (Henriksson *et al*. 1990). The filters were then exposed to X-ray films (XAR, Kodak) at various times to achieve optimal exposures for each individual sample. Some nylon filters were used for several sequential hybridizations with different probes. The previously used probe was removed and the successful removal of the old probe was examined by exposure of stripped filters to X-ray films as described (Henriksson *et al*. 1995).

Chromosome size estimation

The sizes of the chromosomes were estimated from the central, most dense, part of either ethidium bromide-stained, size-standard chromosomes or hybridization spots revealed by the used markers. For each PFGE gel a standard curve was constructed based on the migration of the size standards. The sizes of hybridizing chromosomes were divided into 5 kb size intervals which were used for aCSDI

Table 1. *Trypanosoma cruzi* samples

(Classification: [a]zymodeme according to Ready & Miles (1980); [b]clonet according to Tibayrenc *et al*. (1986); [c]recommended nomenclature (Momen, 1999) with clade classification according to Barnabé *et al*. (2000) and Brisse *et al*. (2000, 2001). Between parentheses: code used in the present study.)

T. cruzi sample	Geographical origin	Host origin	Zymodeme[a]	Clonet[b]	Lineage[c]
Sylvio X10 cl.7 (X10/7)	Belém, Brazil	Human (Acute case)	ZI-related	N.A.	I
Sylvio X10 cl.1 (X10/1)	Belém, Brazil	Human (Acute case)	ZI	17	I
OPS 22 (OPS22)	Cojedes, Venezuela	*Panstrongylus geniculatus*	ZI-related	21	I
OPS 21 cl.11 (OPS21)	Cojedes, Venezuela	Human	ZI-related	19	I
13379 cl.7 (13379)	Santa Cruz, Bolivia	Human (Acute case)	ZI-related	19	I
P 209 cl.1 (P209)	Sucre, Bolivia	Human (Chronic case)	ZI-related	20	I
P11 cl.3 (P11)	Cochabamba, Bolivia	Human (Chronic case)	ZI-related	20	I
Sp 104 cl.1 (SP104)	Region IV, Chile	*Triatoma spinolai*	ZI-related	19	I
Cutia cl.1 (CUTIA)	Espirito Santo, Brazil	*Dasylprocta aguti*	ZI-related	19	I
Cuica cl.1 (CUICA)	São Paulo, Brazil	*Opossum cuica philander*	ZI-related	20	I
Esquilo cl.1 (ESQUILO)	São Paulo, Brazil	*Sciurus aestuans ingramini*	ZI-related	20	I
Gamba cl.1 (GAMBA)	São Paulo, Brazil	*Didelphis azarae*	ZI-related	19	I
SO 34 cl.4 (SO34)	Potosi, Bolivia	*Triatoma infestans*	ZI-related	20	I
LGN (LGN)	Region IV, Chile	Human (Chronic case)	ZI-related	19	I
CBB cl.3 (CBB)	Region IV, Chile	Human (Chronic case)	ZII-related	33	IIb
Tu 18 cl.2 (TU18)	Tupiza, Bolivia	*Triatoma infestans*	ZII-related	32	IIb
Esmeraldo cl.3 (ESM/3)	Bahia, Brazil	Human	ZII	30	IIb
CAN III cl.1 (CANIII)	Belém, Brazil	Human	ZIII	27	IIa
MN cl.2 (MN)	Region IV, Chile	Human (Chronic case)	ZII-related	39	IId
NR cl.3 (NR)	Region III, Chile	Human (Chronic case)	ZII-related	39	IId
SC 43 cl.1 (SC43)	Santa Cruz, Bolivia	*Triatoma infestans*	ZII-related	39	IId
Bug2149 cl.10 (V2149)	Rio Grande do Sul, Brazil	*Triatoma infestans*	ZII-related	39	IId
Bug2148 cl.1 (V2148)	Rio Grande do Sul, Brazil	*Triatoma infestans*	ZII-related	39	IId
Tula cl.2 (TULA2)	Chile	*Triatoma infestans*	ZII-related	43	IIe
CL Brener (CLBRENER)	Brazil	*Triatoma infestans*	ZII-related	N.A.	IIe
M6241 cl.6 (M6241)	Belém, Brazil	Human (Acute case)	ZII-related	35	IIc

N.A., No available information.

analysis ($aCSDI_5$, Table 2). For the FFAg6 marker, which hybridized to the compression zone of the gel in the b2148 cl.1 and the b2149 cl.1 stocks, a given value of 1800 kb was used for further analysis for these two samples. In addition, size estimation with 50 kb intervals was used for comparing aCSDI and Jaccard's distance ($aCSDI_{50}$ and JD_{50}, respectively). The 50 kb intervals were based on the 5 kb size interval values, e.g. 5 kb size interval values from 525 to 570 were given the value of 550 kb.

Phenetic analyses

Two methods were used for quantifying chromosome size dissimilarity. First, the absolute Chromosomal Size Difference Index (aCSDI) which considers chromosomes as a continuous variable and weight size variation (progressive model): 2 chromosomes showing a 25 kb size difference are considered to diverge less than 2 chromosomes presenting a 200 kb size difference. The method is based on a diploidy hypothesis (Lanar, Levy & Manning, 1981; Tibayrenc, Cariou & Solignac, 1981; Oliveira *et al*. 1999). In the case of 2 stocks each presenting a single chromosomal band (2 co-migrating sister chromosomes), size difference is counted twice. When 1 of the isolates is presenting a double band (2 different-sized homologues), size difference is measured between each of them and the single band of the second stock, and the sum considered for aCSDI. The index was calculated with a basic program developed previously (available on request to J. C. Dujardin) as described elsewhere (Dujardin *et al*. 1995). Secondly, the karyotype divergence between different samples was also calculated by Jaccard's distance method (Jaccard, 1908): $D_{ij} = 1 - a/(a+b+c)$ where $a =$ number of bands that are common to the stocks i and j; $b =$ number of bands present in the first genotype and absent in the second; $c =$ number of bands absent in the first genotype and present in the second. In this procedure, chromosomes are considered as a discrete variable and there is no weighing of size variation (non-progressive model): 2 chromosomes differing by 25 kb are considered to diverge similarly to 2 chromosomes differing by 200 kb. MLEE and RAPD data (Tibayrenc *et al*. 1993; Ben Abderrazak *et al*. 1993) were also processed by Jaccard's distance. UPGMA trees based on the aCSDI and Jaccard's distance data were constructed using the PHYLIP package software (Felsenstein, 1993).

Table 2. Estimated sizes of hybridizing chromosomes into 5 kb size classes

(Numbers refer to estimated chromosome size in kb. *Sample which contained high concentration of DNA which may cause a retardation of the migration in the PFGE gel.)

	Probe						
Tc clone	IF8	CA7.12	CA7.32	Cruzipain	FFAg6	P19	Tc2
X10/7	670, 720	980, 1040	980, 1040	620	780, 850	1050, 1140	1030
X10/1*	660	1040, 1120	1040, 1120	620, 780	800, 860	1080, 1160	1070
OPS22	650	1010	1010	570	720, 830	1320, 1415	1050
OPS21*	780	1030	1030	800	820, 910	1180, 1240	1070
13379	660, 690	1005	1005	570, 740	780, 830	1045	1000
P209	670	960	960	520, 815	880, 945	1015, 1210	990
P11	705	990	990	610, 850	895	1135, 1230	1040
SP104	700	990	990	605, 840	910	1180, 1430	1050
CUTIA	670, 715	980	980	565, 800	880	1080, 1180	980
CUICA	660, 715	980	980	550	835, 910	1035, 1110	1000
ESQUILO	640, 690	980	980	590	850, 920	1055, 1120	990
GAMBA	670, 725	990	990	590	860, 960	1045, 1100	1030
SO34	670, 725	995	995	585	870, 965	1055, 1120	1040
LGN	555, 650	930	930	530, 580	780, 1360	1045	1020
CBB	715, 760	600	630, 690	975, 1090	905, 985	970, 1390	1270, 1365
TU18	820	610	660, 720	1160	960	1090, 1440	1410
ESM/3	740	600	600	760, 1250	900, 955	1020, 1440	1250, 1330
CANIII	1080	590	735, 810	580	1470	1130, 1230	1540
MN	830, 1140	590, 990	890	665, 1170	970, 1460	1055, 1370	1430
NR	770, 1080	580, 940	890	660, 1160	920, 1415	1055, 1350	1380
SC43*	830, 1180	680, 1110	970	760, 1310	990, 1520	1180, 1470	1490
V2149*	830, 1180	680, 1110	950	740, 1240	1030, 1800	1180, 1440	1480
V2148*	830, 1130	680, 1060	950	730, 1240	1030, 1800	1130, 1440	1460
TULA2	810	700, 860	645, 1110	670, 1020	1210, 1340	1230, 1630	1420
CLBRENER	790, 1115	710, 760	590, 1080	625, 1050	1080, 1360	790, 1710	1340, 1380
M6241	1030	680	945, 1010	910, 1060	1310, 1420	1240, 1380	1340

The error of the estimated size of a chromosome was estimated to correspond to 25 kb for each individual hybridizing chromosome (Giannini *et al*. 1990; Dujardin *et al*. 1995). Therefore, a threshold of significance should be considered when interpreting $aCSDI_5$ dendrograms. This was estimated as previously described (Dujardin *et al*. 1995): number of chromosomes studied × 2 (pairs of homologues of a diploid chromosome) × 25 kb, i.e. a threshold value of 300 kb for 6 chromosomes. No threshold was considered for $aCSDI_{50}$ dendrograms, as it was already considered in the 50 kb size class used for it.

Statistical tests

Correlation analysis between phenetic distances calculated from chromosome, MLEE and RAPD data was performed by 2 tests. Linear regression was analysed by the conventional method using the Excel program. Agreement between distances inferred from PFGE on the one hand, MLEE and RAPD on the other hand, was estimated through a non-parametric Mantel's test (Mantel, 1967), with ADE-4 software (Thioulouse *et al*. 1997). Briefly, this test relies on a Monte Carlo simulation with 10^4 iterations, which randomly permutes the different cells of one of the distance matrices. Contrary to the classical correlation test, this randomization procedure does not need any assumptions about the number of degrees of freedom. Correlations were considered to be significant at P values below 0·05.

RESULTS

Chromosome size polymorphism

In order to analyse chromosome size polymorphism within *T. cruzi*, 26 cloned stocks were selected (Table 1). Most of them had previously been phenetically analysed by MLEE and RAPD (Tibayrenc *et al*. 1993). The sample showed a broad geographical and biological diversity; the stocks originated from 4 different countries and represented the 2 major groups *T. cruzi* I and II and the 5 smaller subdivisions of *T. cruzi* II previously identified (Barnabé *et al*. 2000; Brisse *et al*. 2000, 2001).

Chromosomes from the *T. cruzi* stocks were separated by PFGE and hybridized with 7 different markers and the sizes of hybridizing chromosomes were estimated into 5 kb size intervals (Table 2). All 7 markers revealed different hybridization patterns with the exception of CA7.12 and CA7.32 which showed identical patterns in all samples related to *T. cruzi* I and in 1 stock of *T. cruzi* IIb (Esmeraldo). This might be explained by co-migration of 2 non-homologous chromosomes in these stocks. However,

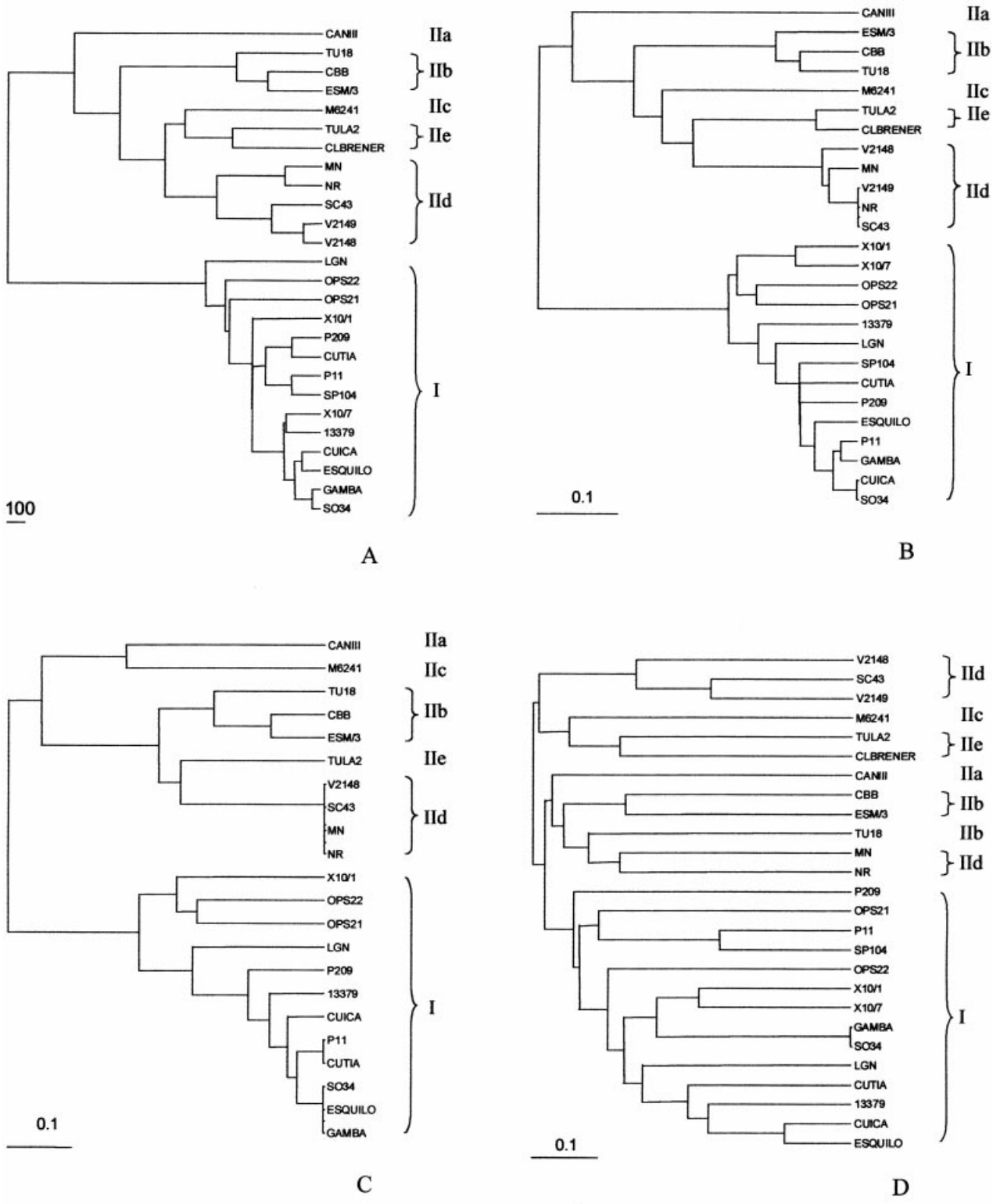

Fig. 1. Phenetic dendrograms constructed with the UPGMA method. The identity of *Trypanosoma cruzi* stocks and the classification in *T. cruzi* I and II and *T. cruzi* IIa–e are indicated in all panels. (A) The $aCSDI_5$ phenetic tree is based on 5 kb size class resolution for 6 different chromosomal markers (marker CA7.32 excluded). The threshold level for each branch is 300 kb. (B) The MLEE dendrogram and (C) the RAPD dendrogram were constructed based on data from Tibayrenc *et al*. (1993) and Ben Abderrazak *et al*. (1993). (D) Jaccard's distance method dendrogram based on 50 kb size classes for 6 different chromosomal markers (marker CA7.32 excluded).

the 2 probes were found to co-hybridize despite size variation of the corresponding chromosomal bands in *T. cruzi* I. It is thus likely that at least in *T. cruzi* I, CA7.12 and CA7.32 are linked on the same chromosome, while they are unlinked in other clades (Henriksson *et al*. 1995).

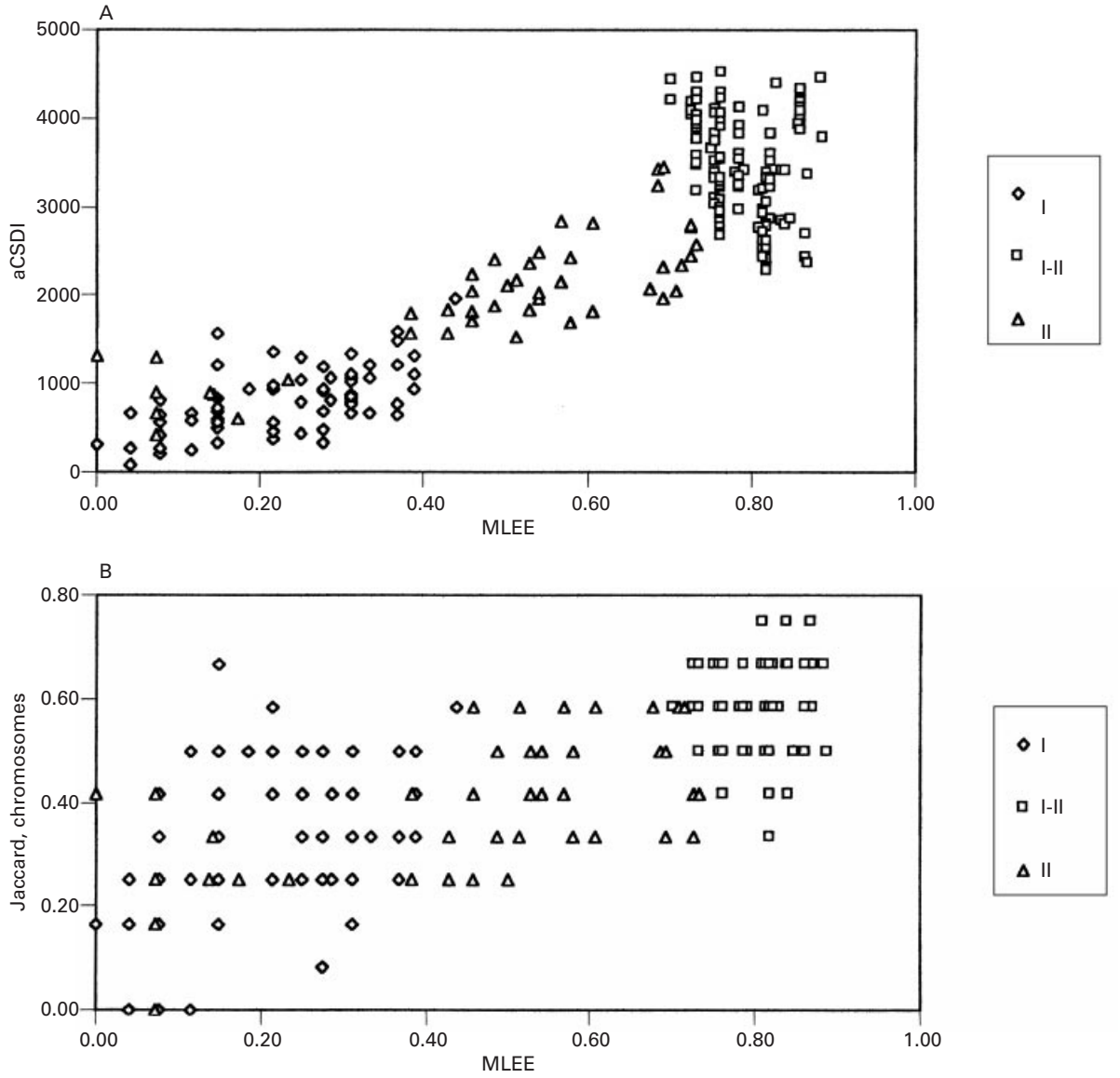

Fig. 2. Relationship between (A) $aCSDI_5$ and Jaccard's distance calculated from MLEE data, and (B) Jaccard's distance calculated from chromosome and MLEE data. Diamond/I, distances calculated within group I; square/I–II, distances between I and II; triangle/II, distances within II.

ACSDI dendrograms from combined chromosome data

The estimated sizes of hybridizing chromosomes divided in 5 kb size intervals (Table 2) were used to perform a phenetic analysis of the different *T. cruzi* stocks under the progressive model ($aCSDI_5$). Because of the possibility of a linkage of CA7.12 and CA7.32 in *T. cruzi* I (and the associated risk of redundancy in aCSDI analyses), combined dendrograms were built up from 6 chromosomes only. No significant differences were observed between combined dendrograms using either CA7.12 or CA7.32 together with the 5 other markers (not shown). The 6 marker-$aCSDI_5$ dendrogram using CA7.12 is shown in Fig. 1A; a 300 kb threshold (see Materials and Methods section) allowed identification of significant clustering. Two major groups were identified, corresponding to *T. cruzi* I and II respectively (Tibayrenc *et al.* 1995). They were separated by a very high aCSDI value (3500 kb). *T. cruzi* I was more homogeneous than *T. cruzi* II, with maximal aCSDI values of 1300 and 2800 kb, respectively. Within *T. cruzi* II, stocks clustered according to each of the 5 subdivisions *T. cruzi* IIa–e.

Comparison of chromosome size polymorphism with MLEE and RAPD data

Simple examination of $aCSDI_5$ dendrogram (Fig. 1A) revealed a similar topology to dendrograms constructed from MLEE and RAPD data (Fig. 1B, C). Clustering was identical and branching of the clusters very similar. The only difference concerned the branching point of *T. cruzi* IIc, close

to (i) *T. cruzi* IIe in the $aCSDI_5$ dendrogram (Fig. 1A), (ii) IId–e in MLEE dendrogram (Fig. 1B), and (iii) IIa in RAPD dendrogram (Fig. 1C). Furthermore, some stocks identical in MLEE and RAPD analyses (*T. cruzi* IId) were very different at chromosome level, with aCSDI values similar to that reported for the whole *T. cruzi* I (Fig. 1A). Phenetic distances calculated on 22 stocks from chromosome ($aCSDI_5$), MLEE and RAPD patterns showed to be significantly correlated (Fig. 2A), with P values of linear regression and Mantel tests lower than 0·01 and high correlation coefficient (0·89 for aCSDI/MLEE and 0·85 for aCSDI/RAPD). Furthermore, scattering of data on the graph was low.

Comparison of numerical methods for chromosome analysis

In order to further validate the progressive model, aCSDI was compared with the Jaccard's distance, a method which does not weigh chromosome size difference; therefore, chromosome size estimated with 50 kb intervals was used. Theoretically, a threshold value should not be applied on this type of dendrogram as the error in size estimation is already considered when defining the 50 kb size classes. The $aCSDI_{50}$ dendrogram was very similar to the $aCSDI_5$ dendrogram (not shown). All stocks, with 3 exceptions (P11, 13379 and LGN), showed conserved location between the $aCSDI_5$ and $aCSDI_{50}$ dendrograms and the maximal aCSDI values between all stocks, and within *T. cruzi* I and II were approximately the same. However, with the Jaccard's distance, much longer branches were observed on the dendrogram (Fig. 1D), which generates statistically uncertain branching points. Furthermore, even if stocks of *T. cruzi* I clustered together, the relationships within *T. cruzi* II were different: (i) 2 clusters were individualized, 1 (containing members of *T. cruzi*, IIa, IIb and IId) being closer to *T. cruzi* I than to other members of *T. cruzi* II, and (ii) stocks of *T. cruzi* IId did not cluster together. Jaccard's distances were significantly correlated with distances calculated from MLEE and RAPD data ($P < 0·01$); however, correlation coefficients were lower (0·75 for Jaccard/MLEE *vs*. 0·89 for aCSDI/MLEE and 0·74 for Jaccard/RAPD *vs*. 0·85 for aCSDI/RAPD). This reflects a much higher dispersion of points on the corresponding distance graphs (see example for Jaccard/MLEE in Fig. 2B), in contrast to the low scattering observed with aCSDI (Fig. 2A).

DISCUSSION

In this study, 2 models were considered to analyse the evolutionary significance of chromosome size polymorphism among 26 stocks representative of *T. cruzi*. Most of the samples had previously been characterized with the widely used MLEE and RAPD (Tibayrenc *et al*. 1993), which made it possible to compare relationships as estimated with these different genetic markers. Fundamental difference between models used for karyotype analysis concerned the progressive (aCSDI; Dujardin *et al*. 1995) or non-progressive (distance of Jaccard, 1908) nature of chromosome size variation.

The progressive model was supported by our results. Firstly, correlation between phenetic distances calculated from chromosome data on one hand, and MLEE/RAPD data on the other hand, were better with aCSDI than with Jaccard's distance. Secondly, in contrast to Jaccard-based dendrograms, aCSDI dendrograms were very similar to those built-up from MLEE and RAPD data: (i) identical structuring in 2 major lineages and 5 lower subdivisions, and (ii) very similar branching. Besides the general agreement between aCSDI and MLEE/RAPD analysis, there were some stocks that showed a higher divergence at chromosome level (e.g. *T. cruzi* IId). The high discriminatory power of molecular karyotyping has been reported previously in *Leishmania* (Dujardin *et al*. 1998) and is likely due to the different nature and frequency of the events involved in chromosome variation (gene rearrangement) and MLEE/RAPD polymorphism (mostly point mutations).

The theoretical model underlying aCSDI presumes that the more the size of homologous chromosomes differs, the more parasite stocks are divergent (Dujardin, 1995). The model fits with the molecular bases of chromosome size variation in Trypanosomatids, which are thought to consist essentially in variation in copy number of tandemly repeated genes (Wagner & So 1990; Campetella *et al*. 1992; Åslund *et al*. 1994; Henriksson *et al*. 1995, 1996; Victoir *et al*. 1995; Inga *et al*. 1998; Kebede *et al*. 1999), a process prone to be progressive. Obviously, chromosome size variation may be bidirectional (amplification or deletion) and reversible, with a consequent high risk of convergence. The congruence here observed suggests that this risk is minimal or not higher than for RAPD and MLEE, certainly when several chromosomes are considered together. The gradual nature of expansion/compression could explain the progressive hierarchization observed within each of the 2 main lineages of *T. cruzi*. However, it hardly explains the long branches existing between the 2 major lineages, and reflecting dramatic size-differences observed between variants of several chromosomes. Corresponding high aCSDI values could reflect an ancient division between both lineages that would have allowed mutations to accumulate, a concept supported by several reports (Tibayrenc *et al*. 1986, 1993; Souto *et al*. 1996; Nunes, de Carvalho & Buck, 1997). The lack of size intermediates between chromosome variants characterizing the 2 subpopulations might eventually be explained by the absence of such intermediates in the present sample or by a natural selection against

them. However, another explanation might be proposed, i.e. major rearrangements like chromosome fusion or splitting. The case of the CA7.12 and 7.32 chromosomes strongly supports this hypothesis, as both markers seemed to be linked in *T. cruzi* I but not in *T. cruzi* II. This suggests the ancient occurence of either a chromosome breakage or a fusion (Henriksson *et al*. 1995), as also reported in *Leishmania* (Britto *et al*. 1998; Wincker *et al*. 1996). The occurrence of such mechanisms would not impede the use of aCSDI; on the contrary, the large size differences associated with chromosome fusion/fission will have a major weight in the analysis, and thus fit with the evolutionary weight of the event. Physical mapping of the corresponding chromosome size variants should allow confirmation of this hypothesis.

Strong agreement between different kinds of genetic markers as reported here is a sign of linkage disequilibrium and constitutes classical circumstantial evidence for clonal evolution, a hypothesis previously proposed to account for the genetic polymorphism of *T. cruzi* (Tibayrenc *et al*. 1986). An alternative hypothesis, although far less parsimonious, would be to infer the presence of 6 biological species within *T. cruzi*. However, even within the lesser subdivisions, clear evidence for rarity of sexual event is present (Barnabé *et al*. 2000). Noteworthy, CL Brener, the stock selected for the *T. cruzi* genome sequencing project was recently shown to be a representation of these rare hybridization events (Machado & Ayala, 2001). Interestingly, this is supported by our present karyotype data, as CL Brener (i) showed 2 different size variants for each of the studied chromosomes, (ii) some of them (like 1F8, CA7.32 and cruzipain) presenting the large size differences mentioned above.

In conclusion, our results showed that despite an extensive size polymorphism, chromosomes constitute valuable markers for evolutionary analyses among Trypanosomatids, when progressive size variation is taken into account. Identification of major evolutionary groups should call for particular attention to the genic content of chromosomes responsible for major structuring, and to their potential relationship with biological (including medical) differences.

We thank Maria Rydåker for excellent technical assistance. This work was supported by financial support from the Swedish Agency for Research Co-operation with Developing Countries (SAREC), the Economical Union (TS3-CT92-0155) and FONDECYT-Chile (193/1043).

REFERENCES

APT, W., AGUILA, X., ARRIBADA, A., GOMEZ, L., MILES, M. A. & WIDMER, G. (1987). Epidemiology of Chagas' disease in northern Chile. Isoenzyme profiles in *Trypanosoma cruzi* from domestic and sylvatic transmission cycles and their association with cardiopathy. *American Journal of Tropical Medicine and Hygiene* **37**, 302–307.

ÅSLUND, L., HENRIKSSON, J., CAMPETELLA, O., FRASCH, A. C., PETTERSSON, U. & CAZZULO, J. J. (1991). The C-terminal extension of the major cysteine proteinase (cruzipain) from *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* **45**, 345–347.

ÅSLUND, L., CARLSSON, L., HENRIKSSON, J., RYDÅKER, M., TORO, G. C., GALANTI, N. & PETTERSSON, U. (1994). A gene family encoding heterogeneous histone H1 proteins in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* **65**, 317–330.

BARNABÉ, C., BRISSE, S. & TIBAYRENC, M. (2000). Population structure and genetic epidemiology of *Trypanosoma cruzi*, the agent of Chagas' disease: a multilocus enzyme electrophoresis approach. *Parasitology* **120**, 513–526.

BEN ABDERRAZAK, S., GUERRINI, F., MATHIEU-DAUDE, F., TRUC, P., NEUBAUER, K., LEWICKA, K., BARNABE, C. & TIBAYRENC, M. (1993). Isozyme electrophoresis for parasite characterization. In *Protocols in Molecular Parasitology* (ed. Walker, J. M.), Vol. 21, pp. 361–382. Humana Press, Totowa, New Jersey, USA.

BRENER, Z. (1982). Recent development in the field of Chagas' disease. *Bulletin of the World Health Organization* **60**, 463–473.

BRISSE, S., BARNABE, C. & TIBAYRENC, M. (2000). Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. *International Journal of Parasitology* **30**, 35–44.

BRISSE, S., VERHOEF, J. & TIBAYRENC, M. (2001). Characterisation of large and small subunit rRNA and mini-exon genes further supports the distinction of six *Trypanosoma cruzi* lineages. *International Journal for Parasitology* (in the Press).

BRITTO, C., RAVEL, C., BASTIEN, P., BLAINEAU, C., PAGÈS, M., DEDET, J. P. & WINCKER, P. (1998). Conserved linkage groups associated with large-scale chromosomal rearrangements between Old World and New World *Leishmania* genomes. *Gene* **222**, 107–117.

CAMPETELLA, O., HENRIKSSON, J., ÅSLUND, L., FRASCH, A. C. C., PETTERSSON, U. & CAZZULO, J. J. (1992). The major cysteine proteinase (cruzipain) from *Trypanosoma cruzi* is encoded by multiple polymorphic tandemly organized genes located on different chromosomes. *Molecular and Biochemical Parasitology* **50**, 225–234.

DIAMOND, L. S. (1968). Improved method for the axenic cultivation of *Entamoeba histolytica* Schaudin 1903 and *E. histolytica*-like amoeba with trypanosomatids. *Journal of Parasitology* **54**, 715–719.

DUJARDIN, J. C., DUJARDIN, J. P., TIBAYRENC, M., TIMPERMAN, G., DE DONCKER, S., JACQUET, D., AREVALO, J., LLANOS-CUENTAS, A., GUERRA, H., BERMUDEZ, H., HAMERS, R. & LE RAY, D. (1995). Karyotype plasticity in neotropical *Leishmania*: an index for measuring genomic distance among *L.* (*V.*) *peruviana* and *L.* (*V.*) *braziliensis* populations. *Parasitology* **110**, 21–30.

DUJARDIN, J. C., BAÑULS, A. L., AREVALO, J., TIBAYRENC, M. & LE RAY, D. (1998). Comparison of chromosomal and isoenzymatic variation in eco-geographical populations of *Leishmania* (*Viannia*) *peruviana*. *Parasitology* **117**, 547–554.

DUJARDIN, J. C., HENRIKSSON, J., VICTOIR, K., BRISSE, S., GAMBOA, D., AREVALO, J. & LE RAY, D. (2000). Genomic rearrangements in Trypanosomatids: an alternative to the 'one gene' evolutionary hypotheses? *Memorias do Instituto Oswaldo Cruz* **95**, 527–534.

DVORAK, J. A., HALL, T. E., CRANE, M., ENGEL, J. C., MCDANIEL, J. P. & URIEGAS, R. (1982). *Trypanosoma cruzi*: Flow cytometric analysis. I Analysis of total DNA/organism by means of mithramycin-induced fluorescence. *Journal of Protozoology* **29**, 430–437.

ENGMAN, D. M., REDDY, L. V., DONELSON, J. E. & KIRCHHOFF, L. V. (1987). *Trypanosoma cruzi* exhibits inter- and intra-strain heterogeneity in molecular karyotype and chromosomal location. *Molecular and Biochemical Parasitology* **22**, 115–123.

FELSENSTEIN, J. (1993) PHYLIP (Phylogeny Inference Package) Version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle, USA.

GIANNINI, S. H., CURRY, S., TESH, R. & VAN DER PLOEG, L. T. H. (1990). Size-conserved chromosomes and stability of molecular karyotype in cloned stocks of *Leishmania major*. *Molecular and Biochemical Parasitology* **39**, 9–22.

GOLDBERG, S. S. & SILVA PEREIRA, A. A. (1983). Enzyme variation among clones of *Trypanosoma cruzi*. *Journal of Parasitology* **69**, 91–96.

GONZALEZ, A., LERNER, T. J., HUECAS, M., SOSA-PINEDA, B., NOGUEIRA, N. & LIZARDI, P. M. (1985). Apparent generation of a segmented mRNA from two separate tandem gene families in *Trypanosoma cruzi*. *Nucleic Acids Research* **13**, 5789–5804.

HENRIKSSON, J., ÅSLUND, L., MACINA, R. A., FRANKE DE CAZZULO, B. M., CAZZULO, J. J., FRASCH, A. C. C. & PETTERSSON, U. (1990). Chromosomal localization of seven cloned antigen genes provides evidence of diploidy and further demonstration of karyotype variability in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* **42**, 213–223.

HENRIKSSON, J., PETTERSSON, U. & SOLARI, A. (1993). *Trypanosoma cruzi*: correlation between karyotype variability and isoenzyme classification. *Experimental Parasitology* **77**, 334–348.

HENRIKSSON, J., PORCEL, B. M., RYDÅKER, M., SABAJ, V., GALANTI, N., RUIZ, A. M., CAZZULO, J. J., FRASCH, A. C. C. & PETTERSSON, U. (1995). Chromosome-specific markers reveal conserved linkage groups in spite of extensive chromosomal size variation in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* **73**, 63–74.

HENRIKSSON, J., ÅSLUND, L. & PETTERSSON, U. (1996). Karyotype variability of *Trypanosoma cruzi*. *Parasitology Today* **12**, 108–114.

IBAÑEZ, C. F., AFFRANCHINO, J. L., MACINA, R. A., REYES, M. B., LEGUIZAMON, S., CAMARGO, M. E., ÅSLUND, L., PETTERSSON, U. & FRASCH, A. C. C. (1988). Multiple *Trypanosoma cruzi* antigens containing tandemly repeated amino acid sequence motifs. *Molecular and Biochemical Parasitology* **30**, 27–34.

INGA, R., de DONCKER, S., GOMEZ, J., LOPEZ, M., GARCIA, R., LE RAY, D., AREVALO, J. & DUJARDIN, J. C. (1998). Relation between variation in copy number of ribosomal RNA encoding genes and size of harbouring chromosomes in *Leishmania* of subgenus *Viannia*. *Molecular and Biochemical Parasitology* **92**, 219–228.

JACCARD, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise de Sciences Naturelles* **44**, 223–270.

KEBEDE, A., DE DONCKER, S., AREVALO, J., LE RAY, D. & DUJARDIN, J. C. (1999). *Leishmania* of subgenus *Viannia*: size polymorphism of chromosomes bearing mini-exon genes in natural populations is due to rearrangement of the mini-exon gene array. *International Journal for Parasitology* **29**, 549–557.

LANAR, D. E., LEVY, L. S. & MANNING, J. E. (1981). Complexity and content of the DNA and RNA in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* **3**, 327–341.

LAURENT, J. P., BARNABE, C., QUESNEY, V., NOEL, S. & TIBAYRENC, M. (1997). Impact of clonal evolution on the biological diversity of *Trypanosoma cruzi*. *Parasitology* **114**, 213–218.

MACHADO, C. A. & AYALA, F. J. (2001). Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. *Proceedings of the National Academy of Sciences, USA* **98**, 7396–7401.

MANTEL, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209–222.

MILES, M. A., LANHAM, S. M., SOUZA, A. A. & DE POVOA, M. (1980). Further enzymatic characters of *Trypanosoma cruzi* and their evaluation for strain identification. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **74**, 221–237.

MILES, M. A., CEDILLOS, R. A., POVOA, M. M., SOUZA, A. A., PRATA, A. & MACEDO, V. (1981). Do radically dissimilar *T. cruzi* strains (zymodemes) cause Venezuelan and Brazilian forms of Chagas' disease? *Lancet* **20**, 1338–1343.

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

NUNES, L. R., de CARVALHO, M. R. & BUCK, G. A. (1997). *Trypanosoma cruzi* strains partition into two groups based on the structure and function of the spliced leader RNA and rRNA gene promoters. *Molecular and Biochemical Parasitology* **86**, 211–224.

OLIVEIRA, R. P., MELO, A. I., MACEDO, A. M., CHIARI, E. & PENA, S. D. (1999). The population structure of *Trypanosoma cruzi*: expanded analysis of 54 strains using eight polymorphic CA-repeat microsatellites. *Memorias do Instituto Oswaldo Cruz* **94** (Suppl. 1), 65–70.

SCHOFIELD, C. J. & DIAS, J. C. (1999). The Southern Cone Initiative against Chagas disease. *Advances in Parasitology* **42**, 1–27.

SOLARI, A., MUNOZ, S., VENEGAS, J., WALLACE, A., AGUILERA, X., APT, W., BRENIERE, S. F. & TIBAYRENC, M. (1992). Characterization of Chilean, Bolivian and Argentinean *Trypanosoma cruzi* populations by restriction endonuclease and isoenzyme analysis. *Experimental Parasitology* **75**, 187–195.

SOUTO, R. P., FERNANDES, O., MACEDO, A. M., CAMPBELL, D. A. & ZINGALES, B. (1996). DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* **83**, 141–152.

THIOULOUSE, J., CHESSEL, D., DOLÉDEC, S. & OLIVIER, J. M. (1997). ADE-4: a multivariate analysis and graphical display software. *Statistics and Computing* **7**, 75–83.

TIBAYRENC, M. (1995). Population genetics of parasitic protozoa and other microorganisms. *Advances in Parasitology* **36**, 47–115.

TIBAYRENC, M. & AYALA, F. J. (1988). Isozyme variability in *Trypanosoma cruzi*, the agent of Chagas' disease: Genetical, taxonomical, and epidemiological significance. *Evolution* **42**, 277–292.

TIBAYRENC, M., CARIOU, M. L. & SOLIGNAC, M. (1981). Interprétation génétique des zymogrammes de flagellés des genres *Trypanosoma* et *Leishmania*. *Comptes Rendus de l'Académie des Sciences*, *Paris* **292**, 623–625.

TIBAYRENC, M., NEUBAUER, K., BARNABÉ, C., GUERRINI, F., SKARECKY, D. & AYALA, F. J. (1993). Genetic characterization of six parasitic protozoa: Parity between random-primer DNA typing and multilocus enzyme electrophoresis. *Proceedings of the National Academy of Sciences*, *USA* **90**, 1335–1339.

TIBAYRENC, M., WARD, P., MOYA, A. & AYALA, F. J. (1986). Natural populations of *Trypanosoma cruzi*, the agent of Chagas' disease, have a complex multiclonal structure. *Proceedings of the National Academy of Sciences*, *USA* **83**, 115–119.

VICTOIR, K., DUJARDIN, J. C., DE DONCKER, S., BARKER, D. C., AREVALO, J., HAMERS, R. & LE RAY, D. (1995). Plasticity of gp63 gene organization in *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *peruviana*. *Parasitology* **111**, 265–273.

WAGNER, W. & SO, M. (1990). Genomic variation of *Trypanosoma cruzi*: Involvement of multicopy genes. *Infection and Immunity* **58**, 3217–3224.

WINCKER, P., RAVEL, C., BLAINEAU, C., PAGÈS, M., JAUFFRET, Y., DEDET, J. P. & BASTIEN, P. (1996). The *Leishmania* genome comprises 36 chromosomes conserved across widely divergent human pathogenic species. *Nucleic Acids Research* **24**, 1688–1694.

WORLD BANK (1993). *World Development Report 1993. Investing in Health*. Oxford University Press, New York.