

# Genetic relationships and spatial genetic structure among clonal stocks of *Trypanosoma cruzi* in Colombia

M. RUIZ-GARCIA†\*, M. MONTILLA‡, S. O. NICHOLLS‡, L. ANGARITA§  
& D. ALVAREZ†

†Laboratorio de Bioquímica, Biología y Genética Molecular de Poblaciones, Unidad de Genética, Departamento de Biología, Facultad de Ciencias, Pontificia Universidad Javeriana, Cra 7 a no. 43–82, Bogotá DC, Colombia, ‡Instituto Nacional de Salud (I.N.S), Laboratorio de Parasitología, Avenida El Dorado, Carrera 50, Zona 6, Bogotá DC, Colombia and §Laboratorio de Bioquímica, Biología y Genética Molecular de Poblaciones, Departamento de Bioquímica, Facultad de Ciencias, Pontificia Universidad Javeriana, Cra 7 a no. 43–82, Bogotá DC., Colombia

Genetic variability in the protozoan causative agent of Chagas' disease, *Trypanosoma cruzi*, has been analysed in some Latin American countries; Brazil, Bolivia, Chile and Paraguay. Although Colombia is a country displaying enormous biological diversity, few studies have been conducted from the perspective of the population genetics of *Trypanosoma cruzi*. This study was carried out using 23 Colombian stocks of this protozoan, analysed for 13 isoenzyme loci. The main population genetic results were: (1) Colombia is one of the distribution areas where *T. cruzi* appears to have the highest genetic variability and heterogeneity in Latin America; (2) the *Pgm* locus was found in fixed heterozygosity, supporting the presence of diploidy in this organism; (3) the absence of segregating genotypes and the absence of Hardy–Weinberg equilibrium support the view of the existence of a clonal structure as claimed by Tibayrenc and Ayala. Nevertheless, one characteristic of clonal structure, that of over-representation of some identical zymodemes in vast areas of varying environmental conditions, was not found in Colombia. In this country, a strong spatial autocorrelation, with a classic structure of regional patches, was observed.

**Keywords:** clonal structure, Colombia, population genetics, *Trypanosoma cruzi*.

## Introduction

*Trypanosoma (Schizotrypanum) cruzi* is the causative agent of Chagas' disease, originally described by the Brazilian researcher Carlos Chagas in 1909. Humans, as well as domestic and wild neotropical mammals, are susceptible to carrying this flagellate protozoan. *Trypanosoma cruzi* is distributed between the 42° North latitude (northern California) and 43° South latitude (northern Argentina and Chile) (Paho, 1994), with the majority of the 18 million humans infected spanning from Mexico to Argentina and Chile.

Studies have been carried out on the population genetics of this species, its population structure, and the evolutionary events involved in its high genetic diversity (Tibayrenc & Ayala, 1991). Some areas of Latin America have been exhaustively examined to determine the genetic characteristics of local populations of

*Trypanosoma cruzi*. Perhaps the most detailed studies of the genetic characteristics of the organism in countries of Latin America are those of Tibayrenc & Ayala (1988), Apt *et al.* (1987) (Chile), Chapman *et al.* (1984) (Paraguay), and Breniere *et al.* (1991) (Bolivia, Perú and Chile). Although Colombia is a country displaying an enormous biological diversity, few studies have been conducted from the perspective of the population genetics of this organism; the only studies being those of Saravia *et al.* (1987), and the inclusion of five Colombian stocks (MC50, MC52, MC53, MC60 and MC61) belonging to one zymodeme (19) in the study of Tibayrenc & Ayala (1988). These authors found that the countries displaying the maximum number of different zymodemes, for the 15 loci analysed, were Brazil (with 16 in 23 stocks) and Bolivia (with 17 in 57 stocks). For the five Colombian stocks studied, one zymodeme was found, which is also prevalent in Chile, Brazil, Venezuela and Bolivia. Moreover, in the paper by Saravia *et al.* (1987), considerable

\*Correspondence. E-mail: mruiz@javercol.javeriana.edu.co

MLEE (Multilocus Enzyme Electrophoresis) genotype diversity was found for several Colombian *T. cruzi* stocks. Our present study shows the genetic variability and heterogeneity in Colombia to be substantially higher than that previously reported, indeed the highest known in America, since we report at least 20 (possibly 21) new zymodemes in the 23 stocks of *Trypanosoma cruzi* examined, using 13 out of the 15 loci analysed by Tibayrenc & Ayala (1988).

There is currently some controversy about the type of reproductive mechanism and the genetic structure of *Trypanosoma cruzi*. Studies of some parasitic microorganisms have revealed that sexual reproduction is normal. For example, the African *Trypanosoma brucei* shows Mendelian segregation and recombination (Daude & Tibayrenc, 1994), and certain yeasts (*Candida albicans*), whose sexual behaviour is unclear, show genotypes for genetic markers which are in agreement with Hardy–Weinberg expectations (Tibayrenc & Ayala, 1991). Various authors have indeed stated that panmixia is a common condition in many protozoan parasites (Baker, 1989). The clonal structure of certain parasitic protozoans, including *T. cruzi*, has nevertheless been demonstrated in: (1) deviations from Hardy–Weinberg expectations, which include fixed heterozygosity and the absence of Mendelian segregants; (2) linkage disequilibrium, which includes absence of recombinants and correlation between different genetic characters; and (3) the over-representation of identical zymodemes over extensive geographical areas of varying climatic and environmental conditions. The present study supports, in some aspects, the existence of a clonal structure in *Trypanosoma cruzi* and disagrees with a sexual reproductive structure, confirming the view of Tibayrenc & Ayala (1991), that a rigorous population genetic analysis clearly proves the existence of a clonal structure in the reproduction of *T. cruzi*.

## Materials and methods

### Populations studied

Samples from people infected with *Trypanosoma cruzi*, and animal samples from wild reservoirs as well as vectors, were collected during 10 years from 1985 to 1995 in different parts of Colombia. Upon culture, isolation and characterization, 23 different stocks of *Trypanosoma cruzi* were established. The origin and location of these stocks and their hosts and vectors are shown in Table 1 and Fig. 1. The areas of Colombia examined correspond to endemic areas of *T. cruzi* covering a range of altitudes, temperatures and environmental conditions. For comparison, two samples of *Crithidia* sp. were analysed along with the 23 Colombian

**Table 1** Identification name, provenance and hosts from which they are obtained of 23 Colombian *Trypanosoma cruzi* stocks analysed here

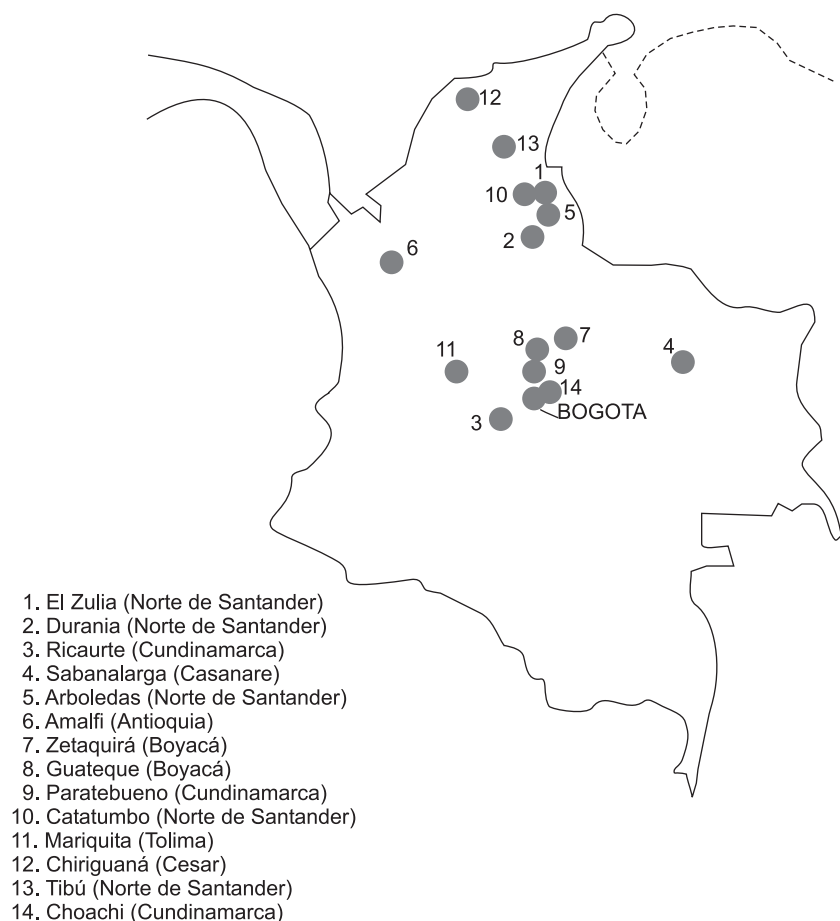
Stock	Provenance	Host/Vector
1	El Zulia (N.S.), San Miguel	<i>D. marsupialis</i>
2	Durania (N.S.), La Troja	<i>D. marsupialis</i>
3	Ricaurte (Cund.), El Callejón	<i>D. marsupialis</i>
4	Sabanalarga (Casanare)	Human
5	Arboledas (N.S.)	<i>D. marsupialis</i>
6	Amalfi (Antioquia)	<i>P. geniculatus</i>
7	Zetaquirá (Boyacá)	Human
8	Guateque (Boyacá)	<i>R. prolixus</i>
9	Paratebuena (Cund.)	Human
10	Catatumbo (N.S.)	<i>R. prolixus</i>
11	Mariquita (Tolima)	<i>D. marsupialis</i>
12	Chiriguaná (Cesar)	Human
13	Tibú (N.S.)	Human
14	Choachi (Cund.)	<i>R. prolixus</i>
15	Ricaurte (Cund.), El Callejón	<i>D. marsupialis</i>
16	Ricaurte (Cund.), El Callejón	<i>D. marsupialis</i>
17	Durania (N.S.), La Troja	<i>D. marsupialis</i>
18	Ricaurte (Cund.), El Callejón	<i>D. marsupialis</i>
19	Sabanalarga (Casanare)	Human
20	Ricaurte (Cund.), El Callejón	<i>D. marsupialis</i>
21	Durania (N.S.), La Troja	<i>D. marsupialis</i>
22	Ricaurte (Cund.), El Callejón	<i>D. marsupialis</i>
23	Tibú (N.S.)	<i>R. prolixus</i>

N.S., Norte de Santander; Cund., Cundinamarca (Colombian Departments). *D. marsupialis*, *Didelphis marsupialis* (Marsupialia); *R. prolixus*, *Rhodnius prolixus*. *P. geniculatus*, *Pastrongilus geniculatus* (insect).

stocks of *T. cruzi*. The stocks were isolated using a modified NNN culture.

### Markers used

For the isoenzymatic analysis,  $10^6$  to  $10^8$  cells/mL from cultured stocks of *T. cruzi* were obtained by centrifugation at 1.000 g for 10 min at 4°C. Upon resuspension in sterile normal saline solution, the cell pellet was kept at –70°C. Electrophoresis in cellulose acetate was carried out according to the procedures of Tibayrenc *et al.* (1985). Thirteen different enzymatic markers were analysed: glucose-6-phosphate isomerase (*Gpi*), malate dehydrogenase (*Mdh*), isocitrate dehydrogenase (*Idh*), phosphoglucumutase (*Pgm*), glutamate dehydrogenase NAD<sup>+</sup> (*Gdh-Nad*<sup>+</sup>), glutamate dehydrogenase NADP<sup>+</sup> (*Gdh-Nadp*<sup>+</sup>), malic enzyme 1 (*Me-1*), malic enzyme 2 (*Me-2*), glucose-6-phosphate dehydrogenase (*G6pdh*), peptidase 1 (*Pep-1*), peptidase 2 (*Pep-2*), 6-phosphoglucuronate dehydrogenase (*6pgdh*) and aspartate aminotransferase (*Got*). The first 12 of these isoenzymes had also been analysed by Tibayrenc & Ayala (1988) in



**Fig. 1** Locality Map of the Colombian *Trypanosoma cruzi* stocks analysed in this study.

their description of 43 American zymodemes, the only difference with our set of polymorphic markers being the use of aspartate aminotransferase in place of leucine aminopeptidase (*Lap*). In their study, these authors had also included aconitase and adenylate kinase, both of which were monomorphic and not included in the profiles of the 43 zymodemes they defined. A comparison between their zymodemes and those presented in this analysis will therefore be possible to a large extent.

#### Population genetic statistical analysis

The population genetic analyses applied were as follows: **1** Hardy–Weinberg Equilibrium. The observed genotypic frequencies were compared with the expected frequencies, assuming Hardy–Weinberg equilibrium using the *F* procedure (Wright, 1965). Because *F* maintains a relationship with the statistic  $\chi^2$ , a significant deviation from zero can be tested:  $\chi^2 = F^2 N(m - 1)$ , with  $m(m - 1)/2$  degrees of freedom, where *N* is the size of the sample, and *m* is the number of alleles at each locus studied (Li & Horvitz, 1953).

**2** Analysis of the degree of diversity and genetic heterogeneity between the 23 stocks of *Trypanosoma cruzi* studied in Colombia. The mean value across loci of heterozygosity was obtained for each stock (Nei, 1978), in order to see whether there was a high variability for this parameter for each of the stocks analysed. The basic statistical measures for a Nei's genetic diversity analysis were obtained (Nei, 1973) for the 23 different stocks.  $H_T$  (genetic diversity in the total Colombian population) was calculated, and therefore were  $H_S$  (the mean genetic diversity in the stocks),  $D_{ST}$  (genetic differentiation between the different stocks which conform to the total Colombian population) and  $G_{ST}$  ( $= D_{ST}/H_T$ , the genetic differentiation between the stocks in relation to the total genetic diversity). High values of  $G_{ST}$  would therefore suggest a high degree of heterogeneity between the different Colombian stocks.

**3** Analysis of the genetic relatedness between the 23 stocks of *Trypanosoma cruzi* studied in Colombia. Two different genetic distances, Nei's (1972, 1978), and Prevosti's (1974), were applied in order to obtain the matrices of genetic distances between pairs of stocks.

The DA distance (Nei *et al.*, 1983) was also calculated for some of the cases. Various types of dendrograms were built based on the genetic distance matrices, using the UPGMA, Single, and WPGMA algorithms. The method of 'neighbour-joining' was additionally applied (Saitou & Nei, 1987). The robustness of the genetic dendrograms obtained was assessed using Felsenstein's bootstrap test (Felsenstein, 1985), which was applied to the dendrogram derived from the UPGMA algorithm and the genetic distance DA (Nei *et al.*, 1983). In order to determine whether the results obtained from the dendrograms remained stable when using other multivariate techniques and for submitting the coordinates obtained to an autocorrelation spatial analysis, a nonmetric multidimensional scaling analysis (MDS) (Kruskal, 1964a,b) was performed. Fifty iterations were run in order to determine the statistic 'final stress'. The 50 iterations could be stopped if a 'minimum stress' of less than 0.001, or a ratio of 'maximum stress' higher than 0.999 were obtained.

4 Autocorrelation analysis of the spatial-structure of the isoenzyme markers studied. Having revealed the extensive genetic heterogeneity for the 13 loci studied in the 23 Colombian stocks of *Trypanosoma cruzi*, we searched for spatial structure of the populations. In

order to obtain a general view of the overall existence of multivariate spatial structure, we applied an analysis of spatial autocorrelation to the coordinates of the first three uncorrelated axes obtained from the 23 stocks, from multidimensional scale analysis (MDS). Each MDS axis I, II and III, represents a different spatial evolutionary event, each affecting all the genetic variables weighted, more or less strongly, and in the sequential order mentioned. This procedure will reduce individual stochastic variation from each of the isoenzyme loci employed.

## Results

The genetic profiles of the 23 Colombian stocks of *Trypanosoma cruzi* studied and those from *Crithidia* are presented in Table 2. The 23 stocks represent 20 different zymodemes, despite *Got* and *Lap* not being included, demonstrating that the genetic variability found in the Colombian stocks is much higher than that reported to date. Zymodeme 19 is that found in all five Colombian stocks in the work of Tibayrenc & Ayala (1988). Here we report 20 new zymodemes, none of which correspond to those previously reported by these authors. In their work Tibayrenc & Ayala (1988) found

**Table 2** Genetic profiles of the 23 Colombian *Trypanosoma cruzi* and two *Crithidia* spp. stocks using 13 isoenzyme loci. The names of the isoenzyme loci codes are shown in the text

Stocks	<i>Gpi</i>	<i>Got</i>	<i>Mdh</i>	<i>Idh</i>	<i>Pgm</i>	<i>GdNad</i>	<i>GdNadp</i>	<i>Me-1</i>	<i>Me-2</i>	<i>G6pdh</i>	<i>Pep-1</i>	<i>Pep-2</i>	<i>6Pgdh</i>
1	7/7	4/4	2/2	2/2	5/7	Absent	3/3	2/2	2/2	2/2	4/4	1/1	—
2	7/7	4/4	2/2	2/2	5/8	2/2	2/2	2/2	2/2	2/2	4/4	2/2	5/5
3	6/6	4/4	3/3	3/3	5/8	Absent	3/3	2/2	2/2	3/3	5/5	2/2	6/6
4	7/7	4/4	2/2	2/2	5/8	2/2	2/2	2/2	2/2	3/3	4/4	Absent	—
5	7/7	4/4	2/2	2/2	5/8	Absent	3/3	3/3	3/3	—	5/5	—	—
6	6/6	2/2	2/2	2/2	5/8	Absent	Absent	3/3	3/3	—	5/5	—	—
7	7/7	4/4	2/2	2/2	5/8	Absent	3/3	3/3	3/3	3/3	4/4	—	—
8	4/4	4/4	2/2	2/2	Absent	Absent	3/3	2/2	2/2	3/3	4/4	1/1	6/6
9	3/3	4/4	2/2	2/2	5/8	Absent	3/3	2/2	5/5	3/3	4/4	2/2	—
10	3/3	2/2	2/2	2/2	5/8	Absent	3/3	2/2	2/2	2/2	6/6	2/2	6/6
11	7/7	4/4	2/2	2/2	5/8	4/4	3/3	2/2	2/2	3/3	2/2	Absent	6/6
12	7/7	7/7	2/2	2/2	5/8	5/5	3/3	3/3	5/5	4/4	5/5	2/2	6/6
13	7/7	4/4	2/2	2/2	5/8	—	3/3	2/2	2/2	3/3	5/5	2/2	6/6
14	6/6	4/4	3/3	2/2	3/4	—	3/3	Absent	3/3	3/3	6/6	Absent	6/6
15	7/7	4/4	2/2	2/2	5/8	2/2	Absent	2/2	2/2	3/3	5/5	2/2	6/6
16	6/6	4/4	3/3	3/3	5/7	Absent	3/3	2/2	2/2	3/3	5/5	2/2	6/6
17	7/7	4/4	2/2	2/2	5/8	2/2	3/3	2/2	2/2	3/3	4/4	2/2	—
18	7/7	4/4	2/2	2/2	5/8	3/3	3/3	2/2	2/2	3/3	5/5	2/2	—
19	7/7	4/4	2/2	2/2	5/8	Absent	3/3	2/2	2/2	3/3	4/4	1/1	6/6
20	7/7	4/4	2/2	2/2	5/8	Absent	3/3	2/2	2/2	3/3	4/4	1/1	6/6
21	7/7	4/4	2/2	2/2	5/8	3/3	3/3	3/3	3/3	3/3	4/4	2/2	—
22	3/3	4/4	2/2	2/2	5/8	3/3	3/3	2/2	2/2	3/3	5/5	2/2	6/6
23	3/3	4/4	2/2	2/2	5/8	3/3	3/3	2/2	2/2	3/3	4/4	1/1	6/6
<i>Crithidia</i>													
1 and 2	2/2	Absent	1/1	2/2	1/1	Absent	Absent	Absent	1/1	1/1	2/2	3/3	1/1

that Latin American countries with high numbers of different zymodemes were Bolivia with 17, Brazil with 16, French Guiana with eight and Chile with six, in contrast to Colombia with one zymodeme. We now show that the genetic variability found in this northern part of South America, with 20 different zymodemes, is as high as, or higher than, that reported in more southern latitudes of the continent. The 20 zymodemes reported in this work are new, and different from the 43 zymodemes reported by Tibayrenc & Ayala (1988) in America. Their profiles are also different from those of the 72 stocks studied in Bolivia, Chile and Peru by Breniere *et al.* (1991), where three different groups of zymodemes were found. The detailed comparison for each locus, contrasting our results with those reported by Saravia *et al.* (1987), Tibayrenc & Ayala (1988) and several new zymodemes recently discovered in Colombia by Márquez *et al.* (1998), is as follows:

**1 *Gpi*:** the homozygote 5/5 had been detected in Colombia by Tibayrenc & Ayala (1988). Saravia *et al.* (1987) and, more recently Márquez *et al.* (1998) detected the genotypes 8/8, 4/7, 4/4, 3/9 and 9/9. In this study we show four different genotypes, of which three had been reported in other locations of Latin America (3/3, 4/4 and 6/6), and the remaining (7/7) is new.

**2 *Mdh*:** the homozygote 2/2 had previously been reported in Colombia, and was seen in many of our stocks. The homozygous form 3/3 is, however, new, both for Colombia and for Latin America.

**3 *Pgm*:** three genotypes new both for Colombia and for Latin America were detected (3/4, 5/7 and 5/8). The homozygotes 1/1, 3/3, 4/4 and 5/5 and the heterozygote 2/3 were already known in Colombia.

**4 *Idh*:** two new genotypes for Colombia were found (2/2 and 3/3), with 3/3 also being new in Latin America. The form 1/1 had previously been reported for this country.

**5 *Gdh-Nad*<sup>+</sup>:** five different genotypes were present, of which 3/3 had already been reported for Colombia. Genotypes 4/4 and 5/5 are reported for the first time for Latin America. A new probable null allele genotype was also detected.

**6 *Gdh-Nadp*<sup>+</sup>:** three genotypes were detected, of which one (2/2) had been detected in Colombia, while another (3/3) had been found in Latin America. A third genotype probably represents a null allele homozygote.

**7 *Me-1*:** three genotypes were found, of which one (2/2) had already been reported in Colombia. Genotypes 2/2 and 3/3 had also been found in other parts of Latin America. A third genotype probably represents a null allele homozygote.

**8 *Me-2*:** three genotypes were found (2/2, 3/3 and 5/5), all new for Colombia, where only one genotype (4/4) had been previously reported. In other parts of Latin

America, genotypes 3/3 and 5/5 had already been reported, but the other genotype (2/2) is apparently new.

**9 *G6pdh*:** four genotypes were detected in Colombia (0/0, 2/2, 3/3 and 4/4). Only three genotypes (5/5, 3/3 and 2/2) had been previously reported for this country, but not the 0/0 and 4/4 genotypes mentioned; therefore, these genotypes are new to Colombia. All of them had already been detected in other parts of Latin America, except for that (0/0) composed of null alleles.

**10 *Pep-1*:** four different genotypes were detected (2/2, 4/4, 5/5 and 6/6) in Colombia. Several genotypes (1/1, 5/5, 4/4 and 2/2) had previously been reported in Colombia. The genotype 6/6, which we report here, is new. All had already been reported for other places in Latin America. Curiously, Márquez *et al.* (1998) detected a large number of heterozygote genotypes at this locus (2/7, 2/5, 1/3, 2/6, 1/7 and 2/3) for Colombia, whereas we did not detect any.

**11 *Pep-2*:** genotypes 1/1 and 2/2 were detected, as well as a null form. Of these, genotype 1/1 had previously been detected in Colombia.

**12 *6pgdh*:** one genotype 4/4 had been reported in the previous study carried out by Tibayrenc & Ayala (1988). In contrast, we report several genotypes (0/0, 5/5 and 6/6) new for Colombia, although previously reported in Latin America, including the null genotype.

In the current work, we compared the 12 isoenzyme markers employed by us, and which were identical to those surveyed by Saravia *et al.* (1987) and Tibayrenc & Ayala (1988). The other isoenzyme, used by us, but not employed by other authors (aspartate aminotransferase), was not introduced in the comparisons.

### Hardy–Weinberg equilibrium

The absence of Hardy–Weinberg equilibrium is significant, and evident for each of the markers used. This applies to all of the stocks considered together. When applying the *F* statistic of Wright, fixation values were obtained revealing the absence of heterozygotes at 12 loci, except for the locus *Pgm*. In this case, applying Wright's *F*, a nonsignificant excess of heterozygosity was obtained ( $F = -0.3158$ ;  $\chi^2 = 14.958$ , 21 d.f.,  $P = 0.825$ ). The absence of Hardy–Weinberg equilibrium of this magnitude is typical of a clonal structure.

### Degree of genetic diversity and heterogeneity between the 23 stocks studied

All the stocks analysed had a similar degree of mean heterozygosity,  $H = 0.0385 \pm 0.0385$ , except for the two stocks of *Crithidia* (from Charalá, Santander), and stock 9 of *T. cruzi* (from Guateque, Boyacá) which had a mean heterozygosity of zero, having no locus fixed in a

**Table 3** Genetic heterogeneity for 13 isoenzyme loci in 23 Colombian *Trypanosoma cruzi* stocks

Loci	$H_T$	$H_S$	$G_{ST}$
<i>Gpi</i>	0.6272	0.0000	1.0000
<i>Got</i>	0.3456	0.0000	1.0000
<i>Mdh</i>	0.4224	0.0000	1.0000
<i>Idh</i>	0.1472	0.0000	1.0000
<i>Pgm</i>	0.6688	0.4400	0.3421
<i>Gdnad</i>	0.5856	0.0000	1.0000
<i>Gdnadp</i>	0.3936	0.0000	1.0000
<i>Me-1</i>	0.4736	0.0000	1.0000
<i>Me-2</i>	0.5376	0.0000	1.0000
<i>G6pdh</i>	0.5088	0.0000	1.0000
<i>Pep-1</i>	0.6560	0.0000	1.0000
<i>Pep-2</i>	0.6656	0.0000	1.0000
<i>6pgdh</i>	0.5920	0.0000	1.0000
All loci	0.5920	0.0338	0.9335

$H_T$ , total genic diversity;  $H_S$ , mean genic diversity in subpopulations;  $G_{ST}$ , genic diversity between subpopulations with regard to the total genic diversity.

heterozygous form. All the other stocks show some heterozygous genotype for the locus *Pgm*. When analysing the genetic heterogeneity of the 13 loci studied, we observed features characteristic of a clonal structure (Table 3). The total genetic diversity ( $H_T$ ) is different for the various loci. The mean genetic diversity ( $H_S$ ) of the stocks for all the markers, except *Pgm*, was zero. This means that all statistics of genetic differentiation between the stocks,  $G_{ST}$ , save that for *Pgm*, had values of 1, which is the maximum possible value. *Pgm* is the only locus with a  $G_{ST}$  value smaller than 1 ( $G_{ST} = 0.342$ ), and shows significant heterogeneity. An average  $G_{ST}$  value of 0.934 was obtained for the 13 loci studied. This represents one of the highest values ever reported for an organism.

#### Genetic relatedness between the stocks of *Trypanosoma cruzi* studied

The genetic distances found between pairs of Colombian stocks of *T. cruzi* are enormous in many cases, suggesting an extensive degree of genetic differentiation between pairs of stocks. Tibayrenc *et al.* (1986a,b) found, for Latin America, values of Nei's distance that ranged from 0.017 to 2.015, with a mean value of  $0.757 \pm 0.478$ . The geographical span we analysed in Colombia is significantly smaller, but the values that we find similarly range from 0.0–0.147 to 1.833, with a mean value similar to that produced for all of Latin America by Tibayrenc & Ayala (1988). This means that the enormous divergence between the stocks studied in Colombia, is of the same order of magnitude as that

found by Tibayrenc & Ayala (1988) for 121 stocks, which covered not only all of Latin America, but also some populations of the USA. The *Crithidia* stocks differed significantly from those of *T. cruzi* with Nei's distance of 2.545. Remarkably, the values between pairs of stocks of *T. cruzi* and *Crithidia* were in some cases smaller than those between certain pairs of stocks of *T. cruzi*.

In order to establish the genetic relatedness between the 23 stocks analysed, 13 trees were generated, of which one showed the highest cophenetic correlation coefficient (in the order of  $r = 0.91$ ), and one had the highest bootstrap values (Fig. 2). The first tree was obtained using the UPGMA algorithm and Nei's genetic distance. In all of the trees constructed the two stocks of *Crithidia* sp. (Charalá, Santander) clearly diverged from the other stocks. In the main cluster there are two groups, one clustering stocks from North-Santander, Boyacá, Casanare, Cundinamarca and Tolima, and a smaller group including stocks from North-Santander, Boyacá, Antioquia and Cesar. The first grouping from the major cluster contains three stocks from Cundinamarca. All the remaining stocks are distributed independently of the Department of Colombia from where they were sampled.

Analyses using other algorithms and genetic distances produced similar results, with only small differences in the relationships between stocks genetically very proximate. The UPGMA tree with the DA distance had the most favourable bootstrap analysis. Those clusters with a bootstrap value higher than 85% were considered to be significant. These were the cluster of the stocks of *Crithidia* vs. the stocks of *T. cruzi* (100%); the split between the main clusters and the first branch separating from all others the stocks representing the populations of Chiriguaná (Cesar), Arboledas (north of Santander) and Amalfi (Antioquia) with a value of 100%. These stocks come from the more northern areas sampled. Another significant cluster is comprised of two stocks from Ricaurte, Cundinamarca (96%). Finally, a significant cluster is found to include two stocks (Sabanalarga (Casanare) and Ricaurte (Cundinamarca)) with 88%. All the other clusters have lower percentage bootstrap values, and are thus not strongly supported.

#### Analysis of the spatial structure

The spatial autocorrelation analysis of the coordinates of the stocks studied for the three main axes derived from the Multidimensional Scaling analysis (MDS), attempts to study simultaneously the spatial information of each variable in each axis of the multidimensional scale. It allows us to better detect global spatial structures (multivariant), reducing the effect of possible



**Table 4** Spatial autocorrelation correlograms using Moran's *I* index of the three principal axes (1, 2 and 3) obtained in a nonmetric multidimensional scaling analysis (MSD) from the Nei's genetic distance matrix and average correlogram for six distance classes from 23 Colombian *Trypanosoma cruzi* stocks

	Distance classes (km)						CgramProb†
	102 km	185 km	267 km	322 km	411 km	580 km	
MSD-1	0.09	-0.39*	-0.10	-0.22	0.12	0.09	<i>P</i> = 0.081
MSD-2	0.12	-0.62**	0.12	0.01	-0.10	0.06	<i>P</i> = 0.000
MSD-3	0.09	-0.10	-0.13	-0.26	-0.21	0.17	<i>P</i> = 0.345
Average	0.10	-0.37*	-0.04	-0.16	-0.06	0.11	

† CgramProb, global correlogram probability.

\**P* < 0.05; \*\**P* < 0.001.

individual differences of the variables involved. With Nei's distance using the Moran's *I* index, it was possible to clearly observe the arrangement of regional patch for MDS I and MDS II (the two most important axes, explaining a large part of the variance), with an approximate size of 160–200 km (mean 185 km). For regional patches it is typical to find that the first distance class is positive while the second distance class is highly negative, whereas the remaining distance classes vary more randomly. The average correlogram for the three MDS axes was highly significant. Using Geary's *c* coefficient, this spatial structure is even more clear (Table 4).

## Discussion

The results of the population genetic analyses presented here support the view of the clonal structure of *T. cruzi* proposed by Tibayrenc (1985), Tibayrenc *et al.* (1986a, b), Tibayrenc & Ayala (1988, 1991) and Tibayrenc *et al.* (1991), as revealed by the following points:

**1** These authors found the high level of polymorphisms for isoenzymes to be relevant to clonal structure. This was the first evidence for clonal structure. They found only one monomorphic locus out of the 16 loci studied (93.3% polymorphism). We found the 13 loci analysed to be polymorphic (100% polymorphism). This is of importance since the number of stocks that we have analysed in Colombia is much smaller than that previously studied by Tibayrenc & Ayala (1988). This would place Colombia as one of the areas where *Trypanosoma cruzi* appears to have the highest genetic variability and heterogeneity. Although a larger number of stocks have been studied in other Latin American countries (Bolivia, Chile and Brazil), the number of new zymodemes defined in Colombia alone (20), is equal to, or higher than, the number of zymodemes described in these countries. Again, when the level of expected heterozygosity is compared, Tibayrenc & Ayala (1988) found a value of  $H = 0.47$  for the whole of America,

whereas we find a higher value  $H = 0.51$  only for Colombia. This could have one of several explanations. If the place with the largest genetic diversity corresponds to the place of original distribution of a species, this would imply that Colombia was the site where the clonal variability was originally established. Alternatively, it could be that a larger number of selective phenomena have affected certain loci, due to the vast diversity of habitats found in the Colombian areas sampled. In addition, due to the complex topography, including the Andes mountains, geographical isolation of various areas sampled is likely to have occurred.

**2** Another argument for clonal structure suggested by these authors was the presence of loci fixed in heterozygous form. They found eight such loci among the 15 loci that they analysed (Tibayrenc & Ayala, 1987). We found one locus in fixed heterozygosity out of the 13 that we analysed, again showing diploidy. Furthermore, in other countries, for example in Paraguay, Chapman *et al.* (1984) and Miles *et al.* (1977), found stocks with homozygous patterns for all the isoenzymes studied, as well as stocks with fixed heterozygous patterns for *Pgm*, *Gpi* and *6pgdh*. Tibayrenc & Ayala (1988) found fixed heterozygous patterns for *Gpi*, *Mdh*, *Idh*, *Me-1*, *Me-2* and *6pgdh* for Latin America as a whole. The locus we detected fixed in heterozygosity in Colombia was *Pgm*, and none of the loci in heterozygosity found by these authors in the rest of Latin America was heterozygous in Colombia. Therefore, although only one locus is fixed in heterozygosity in Colombia, other loci homozygous in this country show fixed heterozygosity in elsewhere of Latin America, indicating an absence or extreme rarity of sexual recombination. As in the findings of Tibayrenc & Ayala (1988) the fixed heterozygosity in the locus *Pgm* in Colombia spans most of the geographical area studied, which is incompatible with meiotic segregation.

**3** No zymodemes which could have been originated by sexual reproduction were found in places from where multiple stocks were sampled, such as was the case in the locations of Ricuarte, Sabanalarga, Durania and Tibú,



in different Departments of Colombia. A similar situation was reported by Tibayrenc *et al.* (1984) in Bolivia, where sympatric zymodemes were not in Hardy–Weinberg equilibrium. We should consider studies on the genetic structure of the vector populations, and how this might influence that of *Trypanosoma cruzi*. Harry *et al.* (1992) demonstrated that *Rhodnius prolixus* (Reduviidae, Triatominae) in Venezuela had a small percentage of polymorphic loci, and low mean heterozygosity (2.5%). The analysis of the Wright's *F* statistics showed that markers such as *Sod* and *Ca* were not in Hardy–Weinberg equilibrium, with an excess of homozygotes ( $F_{IS} = 0.31$  and  $0.26$ , respectively), and that there was a strong genetic heterogeneity for other markers, *Pgi*, *Est-3* and *Ca* ( $F_{ST} = 0.13$ ,  $0.31$  and  $0.15$ , respectively). There was no relation between the genetic distances of pairs of populations of *Rhodnius prolixus* and their geographical distances. This would mean that even if populations of *T. cruzi* with sexual structure could be found, the structure of some of the vectors, at least in parts of Latin America, would affect the genetic structure of *Trypanosoma*, creating an excess of homozygotes, and considerable genetic heterogeneity, although less than that resulting from a clonal structure. It could reinforce the clonal structure, or, alternatively, obscure a sexual structure. Natural selection can move the proportions of genotypes from Hardy–Weinberg expectations, towards both homozygote or heterozygote excess, but is unlikely to affect all markers analysed. However, Dujardin *et al.* (1987) studied the genetic structure of the vector *Triatoma infestans* in Bolivia and found Hardy–Weinberg equilibrium. This implies that it is clonal structure that has created departures from Hardy–Weinberg in *Trypanosoma cruzi*, not the biology of the vector.

4 The over-representation of some identical zymodemes over vast areas of different environmental characteristics has also been considered to be indicative of the existence of clonal structure (Tibayrenc & Ayala, 1988). This is the only characteristic of clonal structure that was not observed in the 23 Colombian stocks of *T. cruzi*. Whereas the clusters found by Tibayrenc & Ayala (1988) in the 43 zymodemes they studied, were geographically randomly distributed, this was not observed for the Colombian stocks, since there is a systematic differentiation between the stocks sampled from the north in the country, and those sampled further south and east. In fact, a good proportion of the correlograms generated (not shown for individual variables) have the classic structure of regional patches, which may mean that, for many alleles, the spatial structure is that typically derived from stochastic processes, such as genetic drift and isolation by distance.

This important difference between our data and those of Tibayrenc & Ayala (1988), could be the result of the different geographical levels involved in the studies. These authors suggest that selection acted on a total general genotype, since some zymodemes (such as 19, 20 and 39) are distributed over wide ecogeographical areas. None of the 20 new zymodemes shown here are found throughout the geographical range analysed. An additional argument by these authors for selection was the consistent isolation of some zymodeme in a given area through time (for example, the zymodeme 34 of Sao Paulo), which might indicate strong local adaptation. This was not found in our stocks. Ricuarte, Sabanalarga, Durania and Tibú, for instance, were sampled several times during the study. Strong differences can be seen between stocks sampled at the same place in the dendrograms in different years, and the stabilizing selection, proposed by Tibayrenc & Ayala (1988), is therefore not observed in the Colombian stocks. Nevertheless, if clonal diversity is considerable it could produce no repeated genotypes, but, otherwise, if the clonal structure is present, a persistent strong linkage disequilibrium should be observed, such as the case in question.

Finally, a comment about the molecular clock. A big fraction of the isoenzymes employed showed correlograms typically describing regional patches, which were caused by stochastic events as mentioned above. However, several less frequent alleles yielded correlograms typical of long-distance patterns (not shown here), which were more likely created by selective events than stochastic ones. Therefore, several less frequent markers affected by selection at long distance could be not used as molecular clocks.

## Acknowledgements

This work was partially supported by Colciencias. Special thanks go to Rosa Chinchilla-Garcia (London, England) for help with English syntax of this manuscript, and to the referees for all their interesting comments and corrections.

## References

- APT, W., AGUILERA, X., ARRIBADA, A., GOMEZ, L., MILES, M. AND WIDMER, G. 1987. Epidemiology of Chagas disease in Northern Chile: Isozyme profiles of *Trypanosoma cruzi* from domestic and sylvatic transmission cycles and their association with cardiopathy. *Am. J. Trop. Med. Hyg.*, **37**, 302–307.
- BAKER, J. R. 1989. Sexual processes in parasitic protozoa. *Int. J. Parasitol.*, **19**, 465–472.
- BRENIERE, S. F., BRAQUEMOND, P., SOLARI, A., AGNESE, J. F. AND TIBAYRENC, M. 1991. An isoenzyme study of naturally

- occurring clones of *Trypanosoma cruzi* isolated from both sides of the West Andes highlands. *Trans. Roy. Soc. Trop. Med. Hyg.*, **85**, 62–66.
- CHAPMAN, M. D., CAFFERY, A., MILES, M. A. AND SWALLOW, D. M. 1984. Enzyme sub-unit numbers in *Trypanosoma cruzi* zymodemes. *Ann. Trop. Med. Parasitol.*, **78**, 541–542.
- DAUDE, F. M. AND TIBAYRENC, M. 1994. Isoenzyme variability of *Trypanosoma brucei*. s. II: Genetic, taxonomic and epidemiological significance. *Exp. Parasitol.*, **78**, 1–19.
- DUJARDIN, J. P., TIBAYRENC, M., VENEGAS, E., MALDONADO, L., DESJEUX, P. AND AYALA, F. J. 1987. Isozyme evidence of lack of speciation between wild and domestic *Triatoma infestans* (Hemiptera, Reduviidae) in Bolivia. *J. Med. Entomol.*, **24**, 40–45.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenetics: an approach using the bootstrap. *Evolution*, **39**, 783–791.
- HARRY, M., MORENO, G. AND GOYFFON, M. 1992. Genetic variability of populations in *Rhodnius prolixus* vector of Chagas' disease in Venezuela. *Evolucion Biológica*, **6**, 175–194.
- KRUSKAL, J. B. 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1–27.
- KRUSKAL, J. B. 1964b. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, **29**, 28–42.
- LI, C. C. AND HORVITZ, D. G. 1953. Some methods of estimating the inbreeding coefficient. *Am. J. Hum. Genet.*, **5**, 107–117.
- MARQUEZ, E., ARCOS-BURGOS, M., TRIANA, O., MORENO, J. AND JARAMILLO, N. 1998. Clonal population structure of Colombian sylvatic *Trypanosoma cruzi*. *J. Parasitol.*, **84**, 184–190.
- MILES, M. A., TOYE, P. J., OSWALD, S. J. AND GODFREY, D. J. 1977. The identification of isoenzyme patterns of two distinct strain-groups of *Trypanosoma cruzi*, circulating independently in a rural area of Brazil. *Trans. R. Soc. Trop. Med. Hyg.*, **71**, 217–225.
- NEI, M. 1972. Genetic distance between populations. *Am. Nat.*, **106**, 283–292.
- NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 3321–3323.
- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- NEI, M., TAJIMA, F. AND TATENO, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. Gene frequency data. *J. Mol. Evol.*, **19**, 153–170.
- PAHO (PAN AMERICAN HEALTH ORGANIZATION) 1994. Chagas disease and the nervous system. *Scientific Publication No. 547*.
- PREVOSTI, A. 1974. La distancia genética entre poblaciones. *Miscellanea Alcobé Publicacions Universitat de Barcelona*, **12**, 109–118.
- SAITOU, N. AND NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
- SARAVIA, N. G., HOLGUIN, A. F., CIBULSKIS, R. E. AND D'ALESSANDRO, A. 1987. Divergent isoenzyme profiles of sylvatic and domiciliary *Trypanosoma cruzi* in the Eastern Plains, piedmont, and highlands of Colombia. *Am. J. Trop. Med. Hyg.*, **36**, 59–69.
- TIBAYRENC, M. 1985. On the microdistribution and sexuality of *Trypanosoma Cruzi*. *Trans. Roy. Soc. Trop. Med. Hyg.*, **79**, 882–883.
- TIBAYRENC, M. AND AYALA, F. J. 1987. *Trypanosoma cruzi* populations: More clonal than sexual. *Parasitol. Today*, **3**, 189–190.
- TIBAYRENC, M. AND 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. AND AYALA, F. J. 1991. Towards a population genetics of microorganisms: the clonal theory of Parasitic Protozoa. *Parasitol. Today*, **7**, 228–232.
- TIBAYRENC, M., ECHALAR, L., DUJARDIN, J. P., POCH, O. AND DESJEUX, P. 1984. The microdistribution of *Trypanosoma cruzi* in southern Bolivia: new isoenzymes profiles and further arguments against Mendelian sexuality. *Trans. R. Soc. Trop. Med. Hyg.*, **78**, 519–525.
- TIBAYRENC, M., CARIOU, M. L., SOLIGNAC, M., DEDET, J. P., POCH, O. AND DESJEUX, P. 1985. New electrophoretic evidence of genetic variation and diploidy in *Trypanosoma cruzi*, the causative agent of Chagas' disease. *Genetica*, **67**, 223–230.
- TIBAYRENC, M., HOFFMAN, A., POCH, O., ECHALAR, L., LE PONT, F., LEMESRE, J. L. ET AL. 1986a. Additional data on *Trypanosoma cruzi* isozymic strains encountered in Bolivian domestic transmission cycles. *Trans. Roy. Soc. Trop. Med. Hyg.*, **80**, 442–447.
- TIBAYRENC, M., WARD, P., MOYA, A. AND AYALA, F. J. 1986b. Natural populations of *Trypanosoma cruzi* and the agent of Chagas disease., Have a complex multiclonal structure. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 115–119.
- TIBAYRENC, M., KJELLBERG, F., ARNAUD, J., OURY, B., BRENIERE, S. F., DARDE, M. L. ET AL. 1991. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 5129–5133.
- WRIGHT, S. 1965. The interpretation of population structure by *F* statistics with special regards to systems of mating. *Evolution*, **19**, 395–420.